UNCLASSIFIED

AD 94409

Armed Services Technical Information Agency

Reproduced by

DOCUMENT SERVICE CENTER

KNOTT BUILDING, DAYTON, 2, OHIO

This document is the property of the United States Government. It is furnished for the duration of the contract and shall be returned when no longer required, or upon recall by ASTIA to the following address: Armed Services Technical Information Agency, Document Service Center, Knott Building, Dayton 2, Ohio.

NOTICE: WHEN GOVERNMENT OR OTHER DRAWINGS, SPECIFICATIONS OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY RELATED GOVERNMENT PROCUREMENT OPERATION, THE U. S. GOVERNMENT THEREBY INCURS NO RESPONSIBILITY, NOR ANY OBLIGATION WHATSOEVER; AND THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED, FURNISHED, OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA IS NOT TO BE REGARDED BY IMPLICATION OR OTHERWISE AS IN ANY MANNER LICENSING THE HOLDER OR ANY OTHER PERSON OR CORPORATION, OR CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

UNCLASSIFIED

AD NO. 24419
ASTIA FILE COPY

FC

ELECTRONICS PERSONNEL RESEARCH

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF SOUTHERN CALIFORNIA

Technical Report No. 13

ELECTRONICS TROUBLE SHOOTING: A BEHAVIORAL ANALYSIS

March 1956

Project Designation NR 153-093

Contract Nonr-228(02)

Principal Investigator

Glenn L. Bryan

Project Director

Nicholas A. Bond, Jr.

Research Assistants

Harold R. LaPorte, Jr.

Lyle S. Hoffman

Prepared for

Personnel and Training Branch
Psychological Sciences Division
Office of Naval Research

DEPARTMENT OF PSYCHOLOGY

UNIVERSITY OF SOUTHERN CALIFORNIA

Reproduction in whole or in part is permitted
for any purpose of the United States Government

PREFACE

This is one of a series of electronics trouble shooting reports by the Electronics Personnel Research group. Titles of other reports in the series are given below.

9. A Methodological Study of Electronics Trouble Shooting Skill: I. Rationale for and Description of the Multiple-Alternative Symbolic Trouble Shooting Test.

   A description of a new type of test format for measuring trouble shooting skill, and a discussion of the conception of trouble shooting on which it is based.

10. A Methodological Study of Electronics Trouble Shooting Skill: II. Intercomparisons of the MASTS Tests, a Job Sample Test, and Ten Reference Tests Administered to Fleet ETs.

    A report of results from two forms of the MASTS Test, its progenitor Job Sample test, and a battery of achievement and ability reference tests.

11. The AUTOMASTS: An Automatically-Recording Test of Electronics Trouble Shooting.

    Mechanical and administrative features of an automatically-recording version of the MASTS Test are described. Test problems, scoring procedures, and a variety of applications are discussed.

12. An Experimental Battery for Measurement of the Proficiency of Electronics Technicians.

    The results of administering the AUTOMASTS Test and a battery of nine printed electronics tests to a large sample of electronics maintenance personnel from the Pacific Fleet.

13. Electronics Trouble Shooting: A Behavioral Analysis.

    A detailed examination of the ways that experienced Electronics Technicians respond in trouble shooting situations.

ACKNOWLEDGMENTS

This research reflects the contribution of a large number of persons within the Military Establishment. Grateful appreciation for general assistance is extended to the Personnel and Training Branch of the Psychological Sciences Division of the Office of Naval Research and the Electronics Coordinator's Section of the Office of Chief of Naval Operations.

This report is an extension of work originated in 1953. At that time, Dr. William W. Grings was the Principal Investigator, Dr. Joseph W. Rigney was the Project Director, and Mr. Stanley A. Summers was a Research Assistant. Much of the planning and data collection for Job Sample and MASTS situations were accomplished by these investigators.

Mr. James R. Ziegler, was responsible for the method (described in Appendix B) for evaluating consistency of trouble shooting approach. He supervised the computer programming and preparation of the reference distributions.

Special thanks are due to the officers and men whose willing cooperation made this work possible.

ABSTRACT

Step-by-step protocols from four data sources are examined with the objective of developing a framework for behavioral analyses of trouble shooting. In an attempt to integrate various facets of such performances, qualitative segments are identified in each protocol, and reference is made to their frequency, position, consistency, and behavioral consequences. Separate treatment is given to special aspects such as redundancy, errors, time, and action rate. Throughout the report, attention is directed to generalizability of the results to different kinds of electronics equipment and to different test formats. The main conclusions are presented in a series of summary statements.

TABLE OF CONTENTS

TABLE OF CONTENTS
(cont'd)

TABLE OF CONTENTS
(cont'd)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

ELECTRONICS PERSONNEL RESEARCH

UNIVERSITY OF SOUTHERN CALIFORNIA

Technical Report No. 13

ERRATA

Page	Correction
20	Footnote 19 should read: "Illustrations of radio IAS categories 1-5 are presented in Appendix A."
41	Delete Footnote 34.
52	Footnote 46 should read: "Negligible relationships also were obtained when the prediagnostic count was limited to checks in the stage where the first component was replaced."
56	Sentence beginning on line 12 should read: "The stages are arranged on the chart from top to bottom according to their order on the signal path."
57	Delete from bottom line: "This figure is derived from Appendix Table G."
58	Delete from two bottom lines: "This figure is derived from Appendix Table H."
63	On line 16, following semicolon, text should read: "the 38 per cent original popularity drops to 24 per cent when all other pulls are considered."

ELECTRONICS TROUBLE SHOOTING: A BEHAVIORAL ANALYSIS

SECTION I. INTRODUCTION

The study of trouble shooting behavior has become increasingly popular. Activities associated with the location and repair of malfunctions have been investigated in electrical (29, 42), mechanical (15, 27), hydraulic (43), electronic (5, 7, 8, 9, 10, 17, 21, 25), and other types of systems. Impetus for these studies has been supplied by the practical urgency of providing effective maintenance for complex military equipment.

But, in addition to contributing toward a better understanding of maintenance knowledge and abilities in order to satisfy immediate military needs, the study of trouble shooting is intrinsically interesting. It represents a complex type of directed thinking in which there are tremendous individual differences. It provides an opportunity for analysts of behavior to extend and develop methods for dealing with long and complex activity sequences. In trouble shooting studies, these behaviors are elicited under circumstances which permit the experimental subjects a larger response repertoire than does the conventional experimental situation. Still, the situation is sufficiently controllable to permit manipulation of relevant variables and the accumulation of manageable data. As a result, step-by-step accounts of trouble

shooting suggest many intriguing possibilities for behavioral analysis. In the repair of electronic equipment, where the underlying functional concepts are often very abstract, these problems of analysis are especially challenging.

The purpose of this report is to present a behavioral analysis of electronics trouble shooting accomplished by experienced Navy technicians. Previous technical reports in this series (3, 5, 20, 21), were devoted to the development of special testing vehicles and scoring methods, with the general goal of providing criteria of technical competence. During the course of our investigations a number of interesting features of trouble shooting came to light. Some of these were not particularly relevant to topics discussed in previous reports. As this body of unreported material grew, the decision was made to write a report discussing various notions about corrective maintenance on the basis of all the information at hand.

In a sense, it is brash to attempt to make general statements about trouble shooting since the term has many meanings and no one knows the limits to which information obtained in one situation may be generalized to others. On the other hand, this research group has collected a large number of detailed sequential records of the activities engaged in by experienced electronics technicians while repairing shipboard electronic gear, while solving problems on electronic laboratory mock-ups, and while attempting to locate and replace defective components in the MASTS and AUTOMASTS tests. It seems reasonable that this body of information should support

statements about trouble shooting, even though they may have to be modified in the light of future developments. It is hoped that such efforts will lead to a more precise specification of a general trouble shooting model.

Organization of the Report

The report consists of seven sections. Section II describes the several types of trouble shooting data used in the analysis. Section III discusses the progressive phases of trouble shooting in relation to a common sequence of trouble shooting events. In Section IV, specific issues involved in trouble shooting are discussed. These are dealt with separately because they do not fit conveniently into the schema presented in the previous section. Section V presents a general analysis of time and rate factors as applied to repair behavior. The sixth section contains a series of summary statements and conclusions concerning the behavior involved in isolating and correcting malfunctions in the experimental situations. Section VII contains a brief discussion of the practical implications of the analysis and raises certain specific points which deserve further examination.

SECTION II. TROUBLE SHOOTING DATA

Descriptions of trouble shooting vary according to the recording system employed, the level of specificity or completeness sought, the emphasis on certain aspects of performance to the exclusion of others, the eventual purposes of the description, and so on. For this reason it is important to state explicitly the

conditions under which the descriptions are made. In the present work, the approach throughout has been empirical. The philosophy has been to evoke and completely record all of the activities engaged in by trouble shooters under circumstances which permitted them the widest possible range of alternative behaviors.

Data from four sources were utilized in the preparation of the present report. As a part of a more comprehensive investigation of the shipboard activities of electronics personnel (22, 23), trained observers kept timed records of the electronics repairs attempted during destroyer training cruises. By prearrangement with the ships' crews, the observers were notified when a casualty occurred. They immediately went to the site of the trouble and remained there until the equipment was restored to satisfactory operation or until repair efforts were abandoned. Observation periods varied from a few minutes to over thirty hours. Continuous records of the activities involved in each repair were made.[1]

A Job Sample trouble shooting test provided a second source of data. A superheterodyne radio receiver and a radar sweep circuit[2] served as the test vehicles. Troubles were introduced into the equipment by substituting faulty parts for good ones, by misaligning stages, and by causing various kinds of concealed discon-

---

[1]However, no attempt was made to produce these records in a standardized form. As a result, their use in this report will be illustrative, rather than supportive.

[2]For the schematic diagrams of these equipments, see Technical Reports 11 and 12 of this series.

tinuities. Each technician taking the test was provided with appropriate manuals, test equipment, hand tools, and a supply of spare parts. Administration was individual, and observers recorded the time, type, and location of each check, adjustment, or replacement made during the trouble shooting attempt.[3]

A third source of data was the MASTS test which employed a device designed to elicit some of the critical behaviors observed in the Job Sample trouble shooting test, and was based on the same radio and radar circuits used in the Job Sample test. The technician taking the test worked from a schematic diagram and other reference materials. At the start of each of 12 problems he was presented with a card showing a trouble symptom for a particular equipment. By lifting corks from holes in a large board, the technician could obtain test information (e.g., voltages, waveforms, resistances, effects of parts replacements) similar to that which he would receive from actual equipment. As in the Job Sample test, administration was individual and timed records were kept of each activity.[4] Inter-observer reliabilities of Job Sample and MASTS data were extremely high. Details concerning the development and administration of both tests have been published elsewhere.

---

[3]Observers had memorized 22 "type of action" categories and several hundred "check points" or locations in the equipments, so that practically all activities could be noted.

[4]MASTS records utilized eight "type of action" categories and the same test points as the Job Sample.

The AUTOMASTS test provided the fourth source of trouble shooting data. This test employed a mechanical, automatically-recording version of the MASTS device. To the technician who takes the test, the machine appears as a black box about the size of a suitcase. By manipulating a set of controls, the technician can obtain different types of information about the functional state of each circuit under the conditions of each problem. Every check or replacement is automatically recorded on paper tape by a printing mechanism inside the box. This feature permits group administration, though every technician has his own testing machine and works independently. Mechanical and procedural features of the AUTOMASTS system are described in previous reports (3, 5).

Altogether, the four data sources represent a natural sequence of research phases. At the outset, the shipboard situation was taken as the ultimate observational setting. As the extended series of shipboard observations indicated, however, systematic research into trouble shooting behavior could be conducted more efficiently under more controlled conditions. The Job Sample test was developed as a trouble shooting task which was complex enough to promote realistic repair attempts and yet simple enough to produce manageable data. Methodologically, the MASTS was an attempt to capture a substantial part of Job Sample variance by means of an abstract representation of electronic equipment. Since this exploratory attempt was relatively successful,[5] the AUTOMASTS

[5]The correlation between Job Sample and MASTS scores were moderately high and positive (21).

version was devised as a means for collecting data under group testing conditions.

There are many obvious differences between looking for troubles in live electronic equipment and solving problems on synthetic devices like the MASTS and AUTOMASTS. On real gear, the perceptual context is much richer; the output symptoms may be more dynamic; and supplementary visual, auditory, and olfactory cues are present. The technician working on real equipment has opportunities to gain information by inspecting front panel indicators or by adjusting control knobs and switches. He also has more opportunity to hook up his test equipment in a faulty manner, to make errors in reading meter scales, to put parts in the wrong place, and to take the same reading in more than one way. He can endanger himself and the equipment by careless work practices.

In the present investigation the effects of such variables were minimized (or eliminated) in order to study the experimental subjects' responses to the functional electronic characteristics of the experimental situation. Within the context of the circuits employed and the problem situations used, it has been shown (21) that the functional electronic nature of the equipment was a primary determiner of the behavior elicited. Undoubtedly, in other situations, using different problems, different types of equipments, and different experimental subjects, there are occasions when some of the features of trouble shooting intentionally ruled out of these studies would be of crucial importance.

However, for the present, it appears more efficient to study behavior elicited in response to the more or less unchanging fundamental principles of electronics, rather than to attempt to study it in relation to the continually changing superficial layout aspects of electronic equipments. Furthermore, it has yet to be demonstrated that the essential, trainable characteristics that distinguish "good" trouble shooters reside in the realm of such factors as perceptual speed, accuracy, sensory acuity, or manual dexterity.

A total of over 1500 trouble shooting records were available for analysis. In this report, the main emphasis will be on results from the Job Sample, MASTS, and AUTOMASTS testings, since these data are in standardized form. Information from the shipboard trouble diaries will be introduced at various places in order to illustrate on-the-job occurrence of certain classes of events exhibited in the more structured records.

In order to make inter-format comparisons, data from the Job Sample and MASTS have been converted to the AUTOMASTS notation system. This means that a considerable amount of Job Sample behavior entries are not analyzed (e.g., the AUTOMASTS notation system has no provision for describing the amount of time the subject spent looking at the schematic, rigging up test equipment, and so on). Even with this conversion, some differences between formats remain. When these differences are important, explicit recognition of the fact will be made in the text. Otherwise, it may be assumed that similar findings were obtained from all formats.

The present report emphasizes results obtained from 81 experienced Naval electronics technicians.[6] These subjects were either petty officers or had been graduated from Class A school; over 50 were Class B school[7] graduates. Practically all had been shipboard ETs for extended cruises. On the basis of technical experience and training, the subject sample was above the general ET average. This sampling bias was consonant with the test development purposes because it seemed that techniques which could discriminate among these highly experienced men would be at least as effective when applied to samples with a greater range of talent.

SECTION III. PROGRESSIVE PHASES OF TROUBLE SHOOTING

In order to facilitate discussion of the sequential trouble shooting procedure, material in this section is organized in accordance with a standardized schema common to many trouble shooting efforts.[8] In this context, trouble shooting consists of a

---

[6] Testing also has been conducted with Aviation Technicians (AT's) and civilians.

[7] Class A is the basic Navy ET course; it consists of 30-42 weeks of training. Class B is a more advanced 28-week course and is usually given after the technician has completed several years service. Since these men were graduated, courses have been shortened.

[8] To discover significant similarities and differences among performances, a standard frame of reference is required. The schema employed in this report represents one effort to develop and utilize a primitive reference system. It is the ultimate hope that eventually adequate "units" of sequential behavior will be delineated so that the more quantitative analyses of all types of sequential behavior will be possible.

series of actions. Each action is either a check (e.g., an activity involved in obtaining information by means of test equipment at some point in the circuit in order to determine the current state of the circuit), or a replacement (substitution of a good tube, resistor, capacitor, etc., for one already in the circuit).

Whenever a technician trouble shoots, he takes some initial action. The discussion in this section begins with an analysis of the initial actions taken by members of the subject group. The analysis relates these actions to other parts of the sequence, the original symptoms, and other trouble shooting performances.

While much can be learned from the initial action alone, a larger segment of the early part of the performance record is required to provide an adequate indication of the method of attack employed by the trouble shooter. It was convenient for the present research to consider the first five actions of each record and to relate those portions of the record to other performance features. The segment of behavior covered by the first five actions is referred to as the initial attack sequence (IAS).

The early portions of most trouble shooting records are characterized by actions intended to locate the trouble (in a gross sense) to a single stage. Sometimes, this type of activity recurs later in the trouble shooting sequence. Whenever it happens, it is called a localizing sequence. The first time it occurs, it is called the initial localizing sequence (ILS).

Following the initial localizing sequence, one usually finds an isolating sequence (IS). This series of actions is an intensive attempt to find the defective part within the suspect stage. In many cases, there are several sets of isolating activity before the first replacement. The extent to which this occurs and the relationships between various patterns of this type of activity and other characteristics of trouble shooting are discussed.

At some point in the trouble shooting schema, a component replacement is made. This is ordinarily thought of as an attempt to substantiate the subject's hypothesis regarding the cause of the pattern of symptoms available to him. For the most part, it probably is just that. In some instances, other explanations seem more plausible. At any rate, it is interesting to examine that particular action in the light of the actions which occur before it. The set of three actions which immediately precede each replacement is referred to in the report as the pre-replacement block. Similarly, all of the actions before the first replacement are considered as a meaningful segment of performance called the prediagnostic behavior.[9]

In one sense, the act of replacing a component is the culmination of one line of the subject's thinking. If the replacement does not correct the malfunction, he must review (and perhaps revise) his conception of the problem and continue his efforts to find the defective component. The first several actions following

---

[9] We are indebted to Damrin and Saupe (10) for this terminology.

an unsuccessful component substitution are analyzed in an attempt to better understand the nature of this behavior. In terms of the present schema, the set of five actions immediately following an unsuccessful replacement is called the <u>post-replacement block</u>.

<u>The Initial Action</u>

At the beginning of the performances analyzed here, the subject is given the gross output of a malfunctioning equipment. On the basis of this information, his past experience with similar symptoms, and his knowledge of the functional relationships embodied in the equipment, he takes his first action. What governs his behavior at this point?

One thing is apparent -- the subjects' first actions are not random. There is far too much correspondence among the initial actions of each man on different problems, and among different men on the same problem, to indicate chance behavior.[10] The question arises, "If not chance, what does determine what an experimental subject will do initially when he is confronted with a problem?"

As one would expect, there is no simple, general answer to this question. There are several factors involved in the choice of an initial action; the circuitry, the problem, the symptom,

---

[10] For example, more than 70 per cent of all of the initial AUTOMASTS actions taken on the radio are accounted for by only 4 specific actions.

the test restrictions (e.g., format), and the man's own style of trouble shooting all enter into the picture.

In the radio problems, initial actions were influenced by the test format. Sixty-five per cent of the radio Job Sample performances began by adjusting volume controls or tuning knobs, while the most popular starting action in the MASTS and AUTOMASTS was to inject a signal at the audio section and to note its effects at the loudspeaker. As indicated in Table 1, almost all of the men who did not begin by injecting signals started by taking a B+ voltage reading at the power supply.

Table 1

Percentage Breakdown of
Initial Actions

	RADIO			RADAR		
	Job Sample N=216	MASTS N=216	AUTOMASTS N=203	Job Sample N=216	MASTS N=216	AUTOMASTS N=224
Adjusts Front Panel Control	65	2	0	43	7	4
Injects Signal	14	65	59	--	--	--
Checks Waveform	--	--	--	43	74	79
Checks B+ Volts	12	20	30	3	10	5
Other	9	13	11	11	9	12

Initial trouble shooting actions on the radar problems differed according to test format in somewhat the same ways as the radio series. In the MASTS and AUTOMASTS formats most of the ETs started out by using their test oscilloscope. However, in the job

Sample format, test scope readings were equal to adjustments in popularity. As shown in Table 1, over 40 per cent of the initial activities consisted of adjusting the monitor scope controls (centering, focus, and intensity) and tuning the oscillators. This tendency to manipulate the front panel controls as an initial trouble shooting action was observed aboard ship also.

It is an interesting diversion to speculate concerning the meaning of this preliminary control manipulation. Many times it doesn't appear to be an attempt to verify an hypothesis, but rather a kind of exploratory behavior intended to produce additional definition of the problem and to aid the production of trouble shooting hypotheses. In some cases, it seems to be just part of the "warm-up" ritual and, in those cases, it is akin to the tire-kicking and door-opening-and-closing behavior exhibited by prospective automobile purchasers. In the problems used in these studies such manipulations weren't particularly helpful as a means for pin-pointing the site of the difficulty.[11]

One wonders why such preliminary control manipulations do not occur in the MASTS and AUTOMASTS radar problems. There are three plausible interpretations for this fact. First, there are no real knobs to twirl in these formats and perhaps a "symbolic knob" does

---

[11] In some systems, however, this type of behavior is of crucial importance. Tucker (45) has recently stated that failure to functionally isolate troubles by manipulating fron panel controls in the K-system bombing and navigational airborne radar is a prime cause of unsuccessful trouble shooting performances on that equipment.

not have the same response potential as a handy physical knob. Second, the time-delay features of the symbolic formats tend to penalize adjustment-type behaviors. Third, the symbolic formats emphasize the functional electronics relationships embodied in the circuitry and in that context some front panel adjustments (which are made on the real gear) appear inefficient or absurd.

One radar problem had an off-centered monitor scope picture as the trouble symptom. On the symbolic formats, 15 per cent of the technicians began this problem by symbolically manipulating the centering controls. While this is a rather small percentage, it is unusually large when compared with problems which have more generalized original symptoms.

There are a few other occasions when the trouble shooting records show that the technicians were responding directly to the original symptoms. The most obvious cases of this sort are where men responded by immediately replacing a component without taking any check readings. This happened infrequently, but in some cases it succeeded in correcting the malfunction. For the most part, however, the subjects had to make a number of checks before they were able to locate and replace the defective component.

Other aspects of the performances can be examined in terms of the way they begin. For instance, the initial actions can be analyzed to determine the extent to which it is possible to predict the eventual success of each performer from these actions alone. As one might expect, records cannot be assigned reliably

to success-failure categories solely on the basis of their initial actions.[12] Furthermore, it is contrary to the spirit of the performances themselves to place too much importance upon any single action.[13]

Before proceeding to the next segment of the trouble shooting schema, brief mention should be made of an interesting series effect.[14] In successive problems there was a progressive reduction in the inter-subject variability of initial actions. Although the same behavioral opportunities existed on every problem, there was a progressive tendency for the technicians to abandon certain initial actions which were unusual by group standards, and to adopt first actions which were more typical of the group. For example, 19 different initial actions were used on the first radio problem. By the fifth problem, only 9 different starting actions were employed, and 75 per cent of all the records for this

---

[12]In isolated instances one or two actions were fairly popular and rather highly predictive of eventual success. The most striking of these was found in the radio data. When the records from all formats are pooled, it is found that 87 per cent of those records which begin by injecting a signal into the grid of the phase splitter terminate with the correct replacement. This high percentage of success may be contrasted with a 65 per cent overall success ratio.

[13]From time to time in this report, single responses will be discussed in isolation. But, it is important to recognize that discrete actions should not be given the status of fundamental behavioral units. Technicians tend to organize their trouble shooting efforts in terms of action sequences rather than a string of discrete activities. It is likely that such sequences are the primordial behavioral units.

[14]Here the term "series effect" refers to an effect which may be directly attributed to the successive administration of the problems.

problem were accounted for by three specific actions.[15] Since the problems were administered under strictly controlled test conditions, this tendency toward standardization over the series must be attributed to learning arising from experience with the problems. On this basis, it appears that learning will occur during the series, even when no attempt is made to "train" the men in desirable trouble shooting practices. The technicians try certain ways of proceeding and then, as a result of hard experience, give them up in favor of empirically derived procedures. It is heartening that, under these circumstances, the commonly arrived at procedures are the ones that the men have been taught and encouraged to use since the beginning of their training. Apparently they try their own methods until they fail a few times, and then they revert to more accepted procedures. Perhaps this experience is a more convincing demonstration of the effectiveness of the recommended procedures than is a statement of faith given from the lecture platform.

The Initial Attack Sequence

Having examined the first action in the performance records in some detail, one might expect to analyze the second action in a similar manner. One also might consider an analysis of the second action as a strict consequence of the first, the third action as

---

[15] There was an increasing tendency over the series for the technicians to begin by injecting signals into the front part of the audio section.

a consequence of the first two, and so on. However, preliminary inspection (later verified by more careful investigation) revealed that such a deterministic view of these trouble shooting processes did not lead to significant results.

The beginning of a trouble shooting performance is markedly different from its later phases. The technician appears to employ a preconceived plan of action.[16] It is interesting to examine the beginning of the records and to investigate the types of behavior exhibited in those segments. Within the present schema these "beginnings" are called initial attack sequences, and they consist of the first five actions of each record.[17]

Initial attack sequences were abstracted from the radio trouble shooting records so that they could be considered independently of subsequent actions or of the success or failure of each performance. On the basis of certain similarities, the sequences were assigned to clusters. The clusters were reviewed and attempts were made to define a behavioral category for each cluster. This process continued through several steps of redefinition until satisfactory categories were obtained. Efforts were made throughout to utilize categories reported by other investigators.

---

[16] On many occasions, ETs volunteered statements of their plan of attack; sometimes this occurred before they began, and sometimes after the performance had terminated.

[17] The selection of five actions was arbitrary. In these cases it proved quite adequate. There is no indication in the data that substantial changes in the categorization of the performances would result from using six or seven actions as the limit for the initial attack sequence.

In particular, the classification systems proposed by Damrin and Saupe (10), Miller *et al.* (36), Glaser (19), and Saupe (41) were investigated. The performances on the radio were categorized before those on the radar because of the relative simplicity of the radio and its ready adaptation to certain theoretical notions of trouble shooting procedure. It was found later that different sets of categories were required for the radar. Therefore, separate treatments of the initial attack sequences are presented for each type of equipment.

Five mutually exclusive categories were developed for the radio initial attack sequences. Their definitions appear below:

IAS Category 1, Half-split. This category is based on the half-split method discussed by Miller, Folley, and Smith (36). The first checking action occurs in the middle of the set,[18] the second checking action occurs in the middle of the remaining defective half, and the third checking action occurs in the middle of the defective quarter.

IAS Category 2, Middle-to-trouble. An orderly checking procedure where the first checking action is in the center of the receiver, and the succeeding ones are in short steps in the direction of the trouble.

IAS Category 3, Loudspeaker-to-antenna. An orderly checking procedure starting at the output end of the receiver and progressing in relatively short steps toward the antenna.

IAS Category 4, Antenna-to-loudspeaker. An orderly checking procedure starting at the antenna and working toward the loudspeaker.

IAS Category 5, Unsystematic. A sequence in which few localizing checks are made. Many sequences assigned to this category resemble what has been called "probability" trouble shooting (17, 36).

---

[18] All designations such as middle, antenna end, loudspeaker end, middle of defective half, etc. were defined by specific test points for each problem.

Cookbook classificatory procedures[19] were devised so that clerks could assign each initial attack sequence to a category. These procedures obviated the usual problems associated with interscorer reliability. A total of 350 records covering the performances of 70 technicians were categorized in this manner. The records used represented performances on each of the three test formats.

Table 2

Relative Popularity of IAS Categories 1 through 5

	Number of Performances Assigned to Category	Percentage of Total Number of Performances
1. Half-split	63	18
2. Middle-to-trouble	61	18
3. Loudspeaker-to-antenna	141	40
4. Antenna-to-loudspeaker	19	5
5. Unsystematic	66	19

The popularity of the first five IAS categories is shown in Table 2. Category 3, Loudspeaker-to-antenna, is clearly the preferred technique. In contrast, the Antenna-to-loudspeaker method

---

[19] Instructions for categorizing the radio IAS's are presented in Appendix A.

is employed in only five per cent of the cases. The other three categories are about equally popular.

If a man starts to work on the first problem of the series in a certain way, will he adopt the same approach of the following problem? To what extent does he tend to "stick with" a single type of attack sequence? Let us turn to the data.

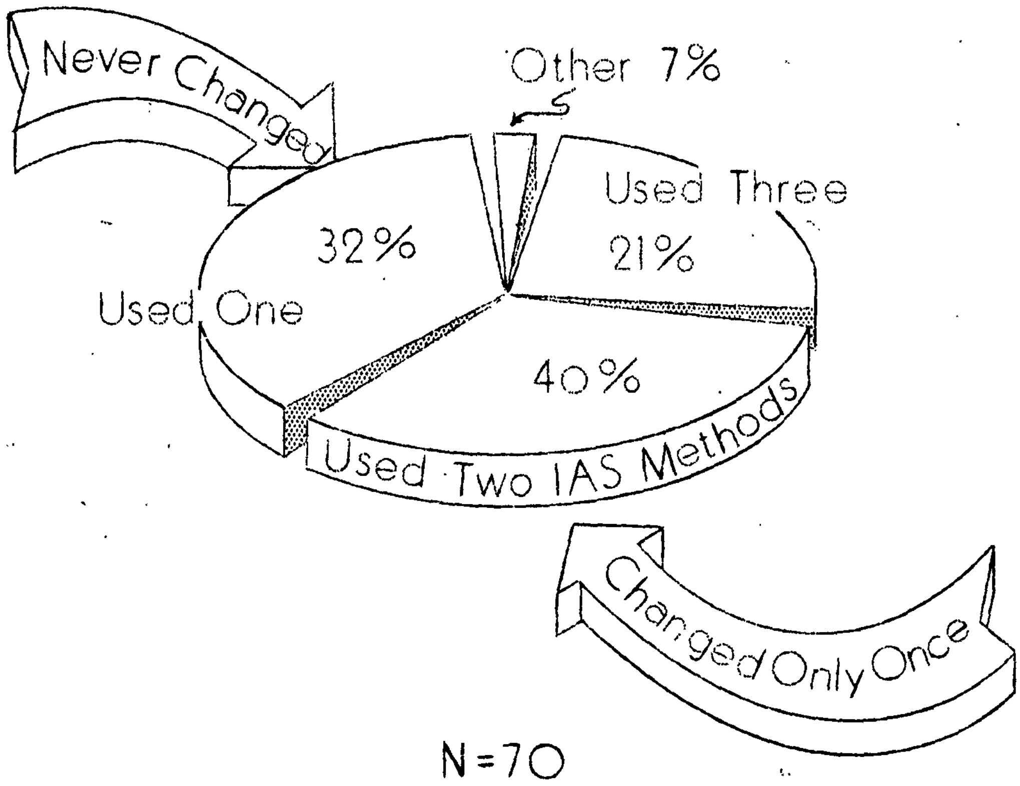


Fig. 1. Constancy of IAS Usage on Radio Problems. The percentage of radio trouble shooters who used a single initial attack sequence throughout the series of five problems and the percentages changing IAS methods one or more times.

Figure I indicates that ETs used relatively few types of IAS when working radio problems. Seventy-two per cent of the group applied only one or two of the five possible types of attack sequence. Almost a third of the group persisted in the same pattern of attack throughout the entire series of radio problems. On the other hand, only three per cent of the group changed their IAS on every problem. Two forms of analysis were employed to specify the technicians' consistency in their use of IAS categories. The first evaluates the departure of obtained results from a random standard. The procedure required the use of a large digital computer and is detailed in Appendix B. Briefly, it consists of producing a statistical distribution of the frequencies with which the IAS methods would have occurred _if_ they had been employed randomly. Except for the element of randomness which was introduced, this distribution was obtained under the same mathematical conditions as the trouble shooting test situation. The mean of the distribution provided by the computer was compared with the mean of the distribution from the experimental data. The experimental mean is almost 13 standard deviations above the randomly-derived mean. On the basis of this comparison it may be confidently asserted that the technicians' consistency in using the IAS methods far exceeded chance expectation.

A second approach to the consistency problem is provided by Ebel-type analysis. This treatment resembles the conventional

internal consistency reliability analysis.[20] The results derived from the radio data are given in Table 3.

Table 3

Inter-problem Consistency of IAS Categories 1 through 5

(N = No. of performers = 70)

Category	Coefficient of Consistency
1. Half-split	.82
2. Middle-to-trouble	.67
3. Loudspeaker-to-antenna	.88
4. Antenna-to-loudspeaker	.67
5. Unsystematic	.68

Table 3 indicates the subjects approached the radio problems with a rather stable "set" toward the five initial attack sequence categories. If a man once adopts a given type of attack, he tends to use it again. Conversely, if he doesn't employ some particular IAS on his first problem or so, he probably will not use it throughout the rest of the series.

[20]However, because of certain dependencies between cell frequencies it is suggested that the coefficients presented here be considered simply as index numbers.

The tendency for the ETs to be consistent is especially significant when one considers the fact that they do not alter their type of attack when a particular IAS results in an unsuccessful attempt to solve the problem. Generally the same approach is tried again anyway.[21] There are several plausible interpretations for this. Perhaps the most reasonable is to assume that they do not attribute their eventual success or failure to the procedure they employ during the very early phases of each trouble shooting effort. As will be seen in Table 4 on page 25, if this is indeed their attitude, they are entirely justified.

This table considers the proportion of performances in each IAS category which terminate in the correct replacement. Although some caution must be exercised in comparing percentages based upon unequal n's, it appears that IAS category and success are generally unrelated.[22] A technician is just about as likely to succeed if he starts in one way as in another.

One also might expect that the particular initial attack sequence employed by the technician would depend upon the type of initial symptom presented. In that case there would be a definite relationship between the type of attack sequence used

---

[21]A chi square value of 0.9, not significant, substantiates this finding.

[22]See Table A of Appendix C.

Table 4

Percentage of the Performances Assigned to IAS Categories 1 through 5 Which Terminated Successfully

Category	Number of Performances in Category	Per Cent Successful
1. Half-split	63	70
2. Middle-to-trouble	61	69
3. Loudspeaker-to-antenna	141	60
4. Antenna-to loudspeaker	19	63
5. Unsystematic	66	43

and the individual problems. In other words, the IAS categories would be "problem-bound." The data in Table 5 on page 26 summarize the relationships between the categories and the problems. It is clear from the table that the five IAS categories occur with about equal frequency on each problem and, therefore, the starting procedures (e.g., the initial attack sequences) are not direct responses to the initial symptoms. The ETs tended to embark upon each problem with a generalized attack sequence that seemed to be determined more by the fact that they were trouble shooting a radio than by the nature of the initial symptom or their success on the preceding problem. We shall see quite a different picture when we come to the discussion of radar problems.

Table 5

The Number of Performances Assigned to IAS Categories 1 through 5 as Related to Five Radio Problems

Problem Number and Cause of Trouble	Half-split	Middle-to-trouble	Loudspeaker-to-antenna	Antenna-to-loudspeaker	Unsystematic
C102. Open Resistor in Detector Filtering Circuit	11	9	30	3	17
C103. Open Resistor Between Detector and AVA	12	7	28	6	17
C104. RF Converter Screen Dropping Resistance Raised	12	14	30	3	11
C105. Detector Filter Capacitor Shorted to Ground	14	12	28	3	13
C106. Detuned Capacitor in Output Tank of IF Amp.	14	19	25	4	8

If, as just shown, the adoption of a particular IAS category is unrelated to success-failure or to special problem conditions (symptoms), then what predictive properties do the categories possess? One promising possibility stems from Miller, Folley and Smith's discussion of the half-split method. As a technical

appendix to the Miller report, Shaycoft formally demonstrated that the half-split method was as least as "efficient" as any other possible method. Efficiency was defined in terms of the number of checking actions required for isolating the defective component. While the IAS phase encompasses only five actions, it appeared reasonable to hypothesize that the half-split performances should have fewer actions. It might also be supposed that the other systematic IAS's should lead to more efficient performances than the unsystematic type of approach. Table 6 summarizes the evidence relevant to comparisons of efficiency.

Table 6

Relationship Between IAS Categories 1 through 5 and Length of Performances on Radio Problems

Category	Average Number of Actions in Performances Assigned to Category
1. Half-split	19.2
2. Middle-to-trouble	22.5
3. Loudspeaker-to-antenna	26.7
4. Antenna-to-loudspeaker	27.8
5. Unsystematic	32.9

As shown in Table 6, the average number of actions in the performances assigned to the Half-split category was smaller than the average number of actions associated with any other category.

An F-test verified that there were significant differences among the categories. Subsequent t-tests were used to evaluate differences between pairs of categories.[23] Further, it was found that the type of IAS could be ordered according to the mean number of moves, and that significant differences in efficiency could be obtained between the categories. Significant differences existed among all except in the case of the "Antenna-to-loudspeaker" IAS. This was not significantly different from either of the other two "systematic" approaches (i.e., the loudspeaker-to-antenna or the middle-to-trouble). In the light of these data it appears that there is a definite relationship between the radio initial attack sequences and the number of actions required to achieve a correct solution.[24]

Before proceeding with discussion of the IAS categories developed for the radar problems it may be useful to review briefly and to evaluate the analytic effectiveness of the IAS categories 1 through 5 as applied to the radio problems. The reader will recall that the first five actions of each trouble shooting record were abstracted so that attention could be focused on the very beginning of the performances which were

---

[23] See Tables B and C of Appendix C

[24] The "efficiency" of these IAS categories, as demonstrated in terms of the length of a trouble shooting attempt, does not hold throughout the performance. Similar efficiency analyses conducted at the point of the first isolation and at the initial replacement did not yield the same ordering of IAS methods. Perhaps the IAS method a man chooses is indicative of his general "attitude" toward efficiency of action. In any case, the exact points where he "saves" actions are difficult to discover.

expected to be more structured than the rest. In a sense, the initial actions are patterns of planned behavior while the rest of the record eventually became highly interactive with the problem situation. As a technician works on a problem he interprets the particularities of each successive action in context and decides on his next action. Consideration of complete records obscures this more highly organized initial segment of behavior. The IAS analysis emphasizes it.

The initial attack sequences for the radio problems are rather highly structured. Furthermore, the structuring is limited to five different categories, each of which can be defined objectively. The most popular approach to the radio problems is signal injection from loudspeaker to antenna. As a group, the technicians tended to select and use one or two of the types of initial approach. They didn't change their method of attack in accordance with the initial symptom or any other problem-related condition. Generally speaking, the initial attack sequences employed were not related to the eventual success or failure of the performers, although significant positive relationships existed between the type of IAS employed and the efficiency of the performance. As will be seen in the following, the results of the radar IAS categorizations were not so clear cut.

An effort was made to assign the radar records to the categories developed for the radio. This appeared to be desirable since these categories then would be generalizable across types of equipment. However, it became apparent that a new set of

categories had to be developed. This was done by the same procedures described for the radio categories. It was not possible to be as objective, however, in the definitions of the new categories. Similarly, it was not possible to develop a set of "cookbook" instructions by which clerks could make the category assignments. Six new IAS categories were developed for the radar performances. Their definitions appear below:

IAS Category 6, Front-to-back. An orderly checking procedure which starts toward the beginning of the sweep-generating pulse path and continues in the direction of the back of the equipment.

IAS Category 7, Back-to-front. An orderly checking procedure which starts toward the rear of the sweep-generating pulse path and continues in the direction of the front of the equipment.

IAS Category 8, Bracket. A procedure characterized by two successive checks at rather widely separated points along the pulse path, with a third check occurring between them. A spatial plot of this type of behavior resembles a "damped oscillation" around a presumed faulty area of the set.

IAS Category 9, Probability. A procedure which suggests that the technician immediately relates the symptom to a specific stage or component and checks his hunch by actions of the type which are ordinarily associated with intra-stage behavior, (e.g., resistance readings or component replacements). It is presumed that probabilistic attack sequences are dependent upon specific previous experiences.

IAS Category 10, Single-stage. An initial attack sequence characterized by a preponderance of generalized checking actions within the limits of a single stage (e.g., taking waveforms).

IAS Category 11, Systemless. An attack characterized by "randomly" distributed checks throughout the set. It is presumably motivated by hope of stumbling on a discrepant reading. Naval trouble shooters call this type of behavior "Easter egging."

In the case of the new IAS categorizations, the procedure was carried out by a group of three judges. No performance was assigned to an IAS category unless at least two of the judges agreed that it belonged there. It was not possible to prescribe extremely objective procedures for assigning the records.[25] A total of 215 performances of 43 ETs working on 10 AUTOMASTS radar problems were dealt with in this manner. The popularity of each of the new IAS categories as applied to radar data is given in Table 7.

Table 7

Relative Popularity of IAS Categories 6 through 11

	Category	Number of Performances Assigned to Category	Percentage of Total Number of Performances
6.	Front-to-back	43	20
7.	Back-to-front	59	27
8.	Bracket	42	20
9.	Probability	37	17
10.	Single-stage	24	11
11.	Systemless	10	5

Four of the categories are just about equally popular. The other two categories, Single-stage and Systemless, were

[25] Illustrations of categorizing the radar IAS's are presented in Appendix A.

used less frequently. Figure 2 shows the extent to which the ETs varied their method of attack for the radar problems.

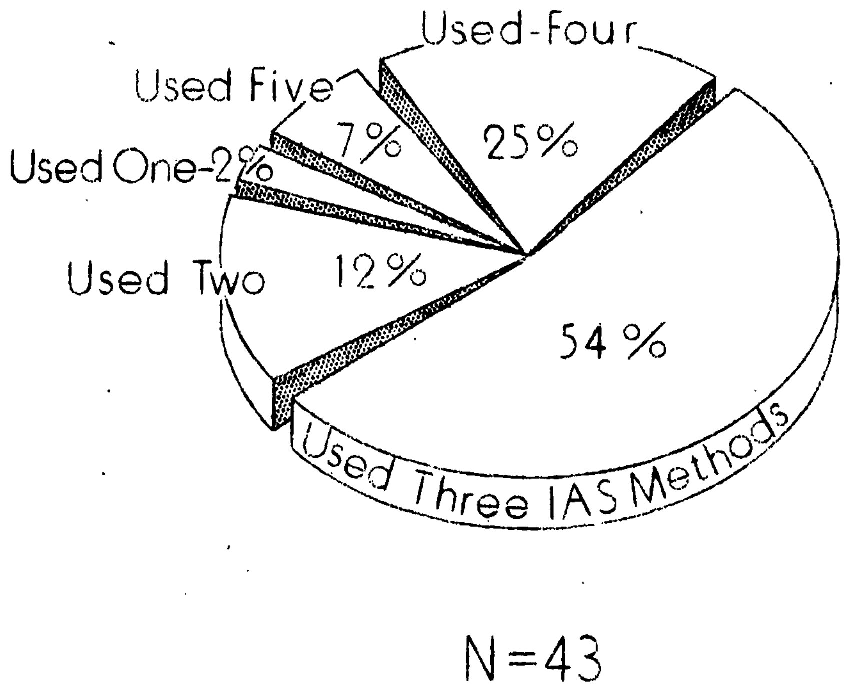


Fig. 2. Constancy of IAS Usage on Radar Problems. The percentage of radar trouble shooters who used a single initial attack sequence throughout the series of five problems and the percentages changing IAS methods one or more times.

In contrast to the tendency of the technicians to employ a small number of different initial attack sequences when working on the radio problems, on the radar problems 86 per cent of the men used three or more different types of initial attack sequence

over the five-problem series. This inconsistency is reflected in the coefficients presented in Table 8.

Table 8

Inter-problem Consistency of IAS Categories 6 through 11

(N = No. of performers = 43)

Category		Coefficient of Consistency
6.	Front-to-back	.50
7.	Back-to-front	.29
8.	Bracket	.19
9.	Probability	.35
10.	Single-stage	.31
11.	Systemless	.04

One is led to wonder why almost all of the ETs demonstrated consistent tendencies to select a single method of attack for the radio problems and to use it more or less throughout the series, but changed their approach from problem to problem in the radar series. One hint is contained in Table 9 on page 34.

Inspection of Table 9 reveals that there was a definite relationship between the initial attack sequence employed and the particular problem being undertaken. For example, half of the occasions where the bracketing procedure occurred were on problem C111. It will be recalled that the radio data did not display this tendency toward "problem-boundedness", but were spread about

Table 9

The Number of Performances Assigned to IAS Categories 6 through 11 as Related to Five Radar Problems

Problem Number and Cause of Trouble	Front-to-back	Back-to-front	Bracket	Probability	Single-stage	Systemless
C110. Open Current Blocking Resistor in Horizontal Positioner	6	4	3	19	7	4
C111. Plate of Sawtooth Generator Tube Open	13	3	21	2	3	1
C112. Leaky Capacitor in Multivibrator Timing Circuit	9	12	5	10	6	1
C114. Shorted Multivibrator Plate Load Resistor	13	9	10	3	5	3
C115. Changed Cathode Bias on Triggered Blocking Oscillator	2	31	3	3	3	1

equally across the problems.[26] Apparently the different functional characteristics of the receiver, as compared with the pulse-forming network, differentially influenced the procedures adopted during the early part of each trouble shooting effort. This may indicate that ETs are less familiar with radar-type equipments than with radio circuits. It also has been suggested that they receive more training on procedures for radio trouble shooting. It may be that the diversity of approaches to radar problems is simply another manifestation of the fact that timing circuits are more difficult to understand fully than are radio receivers.

Since the radar IAS categories could not be objectified, since they were so problem-bound, and since the men were not consistent in their use of the methods further analyses concerning relationships between the methods and success-failure, number of actions, etc., are not undertaken.

However, in order to compare directly the procedures used to attack radio problems with those used for radar problems, all of the data were put into categories 6 through 11. With minor modifications, (e.g., substituting "signal path" for "sweep generating pulse path") the IAS categories 6 through 11 were applicable to the radio problem data. Very little difficulty was encountered in reassigning the radio performances to the new categories. The principal results of this reassignment appear in Table 10 on page 36.

---

[26] See Table 5 on Page 26.

Table 10

Characteristics of Radio Performances When Assigned to IAS Categories 6 through 11

	Category	Number of Performances in Category	Percentage of Total Number of Performances	Coefficient of Consistency
6.	Front-to-back	9	5	(.89)
7.	Back-to-front	81	48	.64
8.	Bracket	54	32	.71
9.	Probability	4	2	(-.11)
10.	Single-stage	16	9	.19
11.	Systemless	6	4	(.23)

When the IAS behavior on the radio problems is compared with the IAS behavior on the radar problems several points deserve comment.[27] The Back-to-front technique is most frequently employed on both types of equipment. The Bracket technique was second in popularity for the radio data, and a close third for the radar. These two categories account for approximately 65 per cent of the total number of records under consideration. With respect to the other categories, considerable inter-equipment

[27]The reader may make these comparisons directly by referring to Tables 7, 8, and 10.

differences may be noted, There were fifteen percentage points difference between the incidence of the front-to-back behavior for the two gears. The same percentage difference occurred with respect to the probability category. In both cases, the differences reflected a greater likelihood for these techniques to be used while working on the radar equipment. Even when the performances are converted to a common category base, the apparent inter-equipment differences with respect to IAS consistency persist. The men were much more inconsistent in their attack upon the radar problems than in their attack on the radio problems. The average coefficient of consistency[28] was .51 for the radio and .29 for the radar problems.

Although the typical technician favored one or two of the IAS methods on the radio problems, there was little indication that he would favor these same methods on the radar problems. As a matter of fact, the average correlation between the use of the various methods on the two types of equipment ($r = -.02$) indicates that the way an ET starts his first radar problem is independent of the way he usually started to work on a radio problem.

To conclude the IAS analysis let us compare the two sets of IAS categories. "Which set is the best?" A direct comparison

---

[28]The averages were computed by z transformation. When the coefficients of consistency were weighted according to the number of cases, the difference becomes more striking (.32 vs. .64).

of the two category sets is possible because both were applied to the radio data.[29]

There are several ways that one may compare the adequacy of category systems. Perhaps the most important is the meaningfulness and interpretability of each. On those grounds there is little to choose between the two sets of IAS classifications under consideration. Both are straightforward and are couched in the trouble shooter's own language. Both are descriptively adequate; both lead to about the same level of understanding of trouble shooting phenomena.

If two category sets are alike in all respects except their level of generality, one would ordinarily prefer the more general. This type of category permits the cross-identification of common behaviors and reveals unique features of the various aspects of the system under study.[30] Categories 6 through 11 are more general than categories 1 through 5, in the sense that all the data can be classified under the former while the latter accommodate only the radio data. On these grounds, 6 through 11 are to be preferred.

The two sets of categories which we have been considering differed according to the degree to which "cookbook" instructions

---

[29]To facilitate the comparison of category sets refer to Tables 2, 3, 4, 6, and 10.

[30]For example, as has been shown, direct comparisons can be made between the types of IAS behavior exhibited on radio and radar problems.

differences may be noted. There were fifteen percentage points difference between the incidence of the front-to-back behavior for the two gears. The same percentage difference occurred with respect to the probability category. In both cases, the differences reflected a greater likelihood for these techniques to be used while working on the radar equipment. Even when the performances are converted to a common category base, the apparent inter-equipment differences with respect to IAS consistency persist. The men were much more inconsistent in their attack upon the radar problems than in their attack on the radio problems. The average coefficient of consistency[28] was .51 for the radio and .29 for the radar problems.

Although the typical technician favored one or two of the IAS methods on the radio problems, there was little indication that he would favor these same methods on the radar problems. As a matter of fact, the average correlation between the use of the various methods on the two types of equipment ($r = -.02$) indicates that the way an ET starts his first radar problem is independent of the way he usually started to work on a radio problem.

To conclude the IAS analysis let us compare the two sets of IAS categories. "Which set is the best?" A direct comparison

---

[28]The averages were computed by z transformation. When the coefficients of consistency were weighted according to the number of cases, the difference becomes more striking (.32 vs. .64).

of the two category sets is possible because both were applied to the radio data.[29]

There are several ways that one may compare the adequacy of category systems. Perhaps the most important is the meaningfulness and interpretability of each. On those grounds there is little to choose between the two sets of IAS classifications under consideration. Both are straightforward and are couched in the trouble shooter's own language. Both are descriptively adequate; both lead to about the same level of understanding of trouble shooting phenomena.

If two category sets are alike in all respects except their level of generality, one would ordinarily prefer the more general. This type of category permits the cross-identification of common behaviors and reveals unique features of the various aspects of the system under study.[30] Categories 6 through 11 are more general than categories 1 through 5, in the sense that all the data can be classified under the former while the latter accommodate only the radio data. On these grounds, 6 through 11 are to be preferred.

The two sets of categories which we have been considering differed according to the degree to which "cookbook" instructions

---

[29] To facilitate the comparison of category sets refer to Tables 2, 3, 4, 6, and 10.

[30] For example, as has been shown, direct comparisons can be made between the types of IAS behavior exhibited on radio and radar problems.

for classifying the materials could be formulated. The set developed for the radio data was strictly objective, and the category assignments could be made by clerks with absolutely no knowledge of electronics. On the other hand, experienced judges were required for the other categorization because completely objective classificatory criteria could not be agreed upon. Therefore, on the basis of inter-judge reliability, Categories 1 through 5 are to be preferred to categories 6 through 11.

Another basis for comparing sets of categories is to consider the "even-ness" with which the cases are spread throughout each set. A set which features evenly spread cases will avoid statistical and interpretative problems arising from varying N's within the categories. Inspection of Tables 2 and 10 reveals that categories 1 through 5 produce a somewhat more even distribution of cases. Three of those categories contain frequencies which approach an ideal spread. On the other hand, four-fifths of the cases fall into only two categories in the other set.

The consistency with which each of the categories was adopted by the technicians during their trouble shooting efforts was discussed earlier. For many analytic purposes, one set of behavioral categories can be considered superior if the coefficients of consistency computed for that set are relatively high.[31]

---

[31] As has been previously indicated (4, 5), inter-problem consistency is not always a necessary or desirable attribute for performance measures.

For categories 1 through 5, the average consistency was .81 as compared with the categories 6 through 11 average of .64.[32] This difference is favorable to categories 1 through 5, though not markedly so.

Finally, a group of comparisons can be made in terms of the discriminability or predictive relevance of the category sets.[33] In order to determine the extent to which the categories discriminated between eventual success, parallel analyses were performed for each category set. Similarly, analyses were carried out to determine the discriminability within sets in terms of length of performance. Neither set shows significant discrimination on the success-failure criterion. The set containing categories 1 through 5 provides significant discrimination in terms of length of performance.

The outcomes of the preceding comparisons between category sets indicate that each possesses advantages and limitations. Categories 1 through 5 are more objective, more consistent, yield a more even distribution of cases, and discriminate in terms of length of performance. The only advantage of categories 6 through 11 is that they are more general.

---

[32] Coefficients were averaged by means of the $\underline{z}$ transformation (13), and weighted according to number of cases represented in each category. The data used in these comparisons are 170 AUTOMASTS radio performances by 34 ETs.

[33] It must be emphasized that IAS categories in both sets were defined independently of other performance characteristics.

The overall conclusions of the IAS analysis can now be restated in summary form. Perhaps the most significant finding is that the first few actions of trouble shooting performances can be classified into a small number of mutually exclusive categories defined in terms of the directional patterns of the checking actions made along a signal path or main pulse line. Very few technicians are completely haphazard at the start of a trouble shooting task. Instead, a coherent, fairly systematic method of attack is adopted. On radio problems, technicians display a marked tendency to stick with a certain search pattern; they are much less consistent on radar problems. A technician's favorite radar starting pattern cannot be predicted from his preferred method on the radio. It is possible, however, to improve predictions of some characteristics of trouble shooting on the basis of the type of initial attack sequence employed.

The Initial Localizing Sequence

The IAS, as discussed above, is limited to the first five actions. In an effort to examine the early part of trouble shooting performances from a different angle, a segment of behavior called the initial localizing sequence (ILS) is defined. This segment consists of all the actions which occur prior to the first isolation to a stage.[34] In contrast with the IAS

---

[34] Criteria for localizing to a stage are described in Appendix A.

concept, the ILS imposes no restriction on the number of actions. Thus an ILS might consist of a very long series of checks which include the five IAS actions, or it even might be shorter than the IAS. In any case, the ILS limits are derived from a logical analysis of the performance phases. It was expected that separate behavioral analyses would be carried out within the ILS definition of starting behavior.

However, investigation revealed that considerable overlap existed between ILS and IAS. The average ILS contained only 7.4 actions.[35] This meant that, on the average, the ILS consisted of the IAS plus two or three additional actions. In view of the ILS-IAS overlap, comprehensive treatment of ILS will not be presented. Only a few findings have been singled out for brief mention.

Since ILS's vary with respect to length, certain relationships between the number of ILS actions and other characteristics can be investigated. The number of ILS actions is fairly consistent over a series of ten problems, as is shown by an Ebel coefficient of .59.[36] Thus, regardless of problem and equipment conditions, a technician is rather consistent with respect to his amount of checking activity before he enters a stage. There is no relationship ($r = -.03$) between ILS length and the total number of problems

[35] Table D of Appendix C gives summary statistics of the ILS for Job Sample, MASTS, and AUTOMASTS performances.

[36] The ILS analyses are based on 340 performances by 34 ETs.

solved. However, the number of ILS actions is significantly related ($r = .46$) to total number of actions.

Table 6 showed there are significant differences among IAS categories (1 through 5) in terms of the length of the average performance in each category. Under those circumstances, the Half-split approach is the most efficient and the Unsystematic approach is the least efficient. This is in accord with expectations. Though one further might expect this order to hold throughout the performances, it does not do so. Indeed, the average performance which starts out with a Category 5 attack (Unsystematic) has a shorter initial localizing sequence than any other type of IAS. The reader will recall that the Unsystematic category was associated with the longest total performances. Other changes occur too; the Loudspeaker-to-antenna attack sequence (which has a moderate number of total actions associated with it) has, by far, the largest number of ILS actions.

The Isolating Sequence

It has been indicated that most of the performances begin with a type of checking activity intended to localize the malfunction to an equipment stage. This generally involves the use of signal-injecting and signal-tracing techniques to discover a break in the signal path (or, in the radar, the pulse path). After seven or eight actions (on the average) the ETs alter their behavior patterns. They restrict their checking activity to a

small section of the equipment and switch to voltage and resistance measurements. In this manner they proceed to check an equipment stage rather intensively with the goal of finding the defective component or rejecting the hypothesis that the faulty part is in the stage under examination. These intra-stage behaviors are termed isolating sequences (IS) within the present schema.

In order to investigate this type of behavior a distinction was made among four types of actions; signal path or pulse path checks, isolating checks, component replacements, and general checks.[37] This classification served as a basis for a set of objective rules for the determination of each isolating sequence.[38]

There is some variation in the number of isolating sequences in different records. The range is from 0 to 19, and the typical performance contains three isolating sequences. Radar problems

[37]A signal path or pulse path check involves the use of such test equipment as signal generators and oscilloscopes to check continuity, determine the presence of a signal or pulse, and to examine its dynamic characteristics. Isolating checks include most voltage measurements, all resistance measurements, and front panel adjustments. Component replacements and most screwdriver adjustments are included in the replacements group. The general checks consist of occasional B+ measurements, filament voltage checks, and incidental AC voltage readings.

[38]Complete instructions for the determination of isolating sequences are contained in Appendix A. Basically, these require that the tabulator consider each action in turn. When an action meets the criteria it is included in the sequence. The sequence continues until an action occurs which does not meet the criteria. Generally speaking, actions of this latter type will be either out-of-stage, signal oriented, checks of the localizing type, or replacements.

have a slightly higher average number of IS's (as compared with the radio problems). The number of IS's observed on a particular problem is highly correlated with the difficulty level of the problem (Rho = .88). The ETs are rather consistent with respect to the number of IS's that they employ on each problem, as shown by an Ebel coefficient of .64. There is no relationship between the number of IS's in a given record and the likelihood that the record will terminate with the correct component replacement (r = -.14). On the other hand, the longer performances generally contain more IS's than the shorter performances. (The correlation between number of IS's and length of performances is .71).

The average isolating sequence contains three actions.[39] It was hypothesized that long IS's were a sign of thoroughness and would be associated with problem success. A Pearson correlation of .46 indicates that there was some tendency in that direction. A similar relationship existed between the average lengths of the isolating sequences and the average lengths of the total records; the coefficient in this case was .50.

The data were examined to determine whether there was a progressive change in the length of the isolating sequences as a function of their order within the records. A correlational

---

[39] This average is based upon 10 AUTOMASTS problems; the range is 1 to 16. It does not include those instances where components were replaced without immediately prior isolating activity in the stage where the component was located.

analysis of successive pairs of IS's indicated that such changes do not occur; i.e., first IS's are about as long as the second ones, and so on.

A procedure was applied to the isolating sequences (IS) to see if action patterns could be discerned within the more limited confines of an equipment stage. Over 1200 isolation sequences were reviewed. Either such action patterns did not exist or were so obscure that they could not be recognized in three-fourths of the cases. Of those patterns that could be reliably classed, the majority were checks between the B+ supply and the plate side of the tube; about twenty per cent of the IS's were of this type. Directional patterns analogous to Front-to-back or Back-to-front were practically non-existent and comprised only about four per cent of the cases. These results provide strong support for the generalization that intra-stage behaviors are relatively disorganized and unsystematic.

Since the typical trouble shooting performance record contains more than one isolating sequence, a number of checking actions commonly occur between the end of one IS and its successor. Although these actions theoretically could be of several types, nearly all of those found in the records were of a generalized checking nature. In many ways, they closely resemble the type of behavior previously referred to as the ILS. They feature the same search for discontinuity at the stage level and the same utilization of signal-injection and signal-tracing techniques, with an occasional power supply check. The average sequence

analysis of successive pairs of IS's indicated that such changes do not occur; i.e., first IS's are about as long as the second ones, and so on.

A procedure was applied to the isolating sequences (IS) to see if action patterns could be discerned within the more limited confines of an equipment stage. Over 1200 isolation sequences were reviewed. Either such action patterns did not exist or were so obscure that they could not be recognized in three-fourths of the cases. Of those patterns that could be reliably classed, the majority were checks between the B+ supply and the plate side of the tube; about twenty per cent of the IS's were of this type. Directional patterns analogous to Front-to-back or Back-to-front were practically non-existent and comprised only about four per cent of the cases. These results provide strong support for the generalization that intra-stage behaviors are relatively disorganized and unsystematic.

Since the typical trouble shooting performance record contains more than one isolating sequence, a number of checking actions commonly occur between the end of one IS and its successor. Although these actions theoretically could be of several types, nearly all of those found in the records were of a generalized checking nature. In many ways, they closely resemble the type of behavior previously referred to as the IIS. They feature the same search for discontinuity at the stage level and the same utilization of signal-injection and signal-tracing techniques, with an occasional power supply check. The average sequence

between IS's consists of four actions. This is considerably shorter than the ILS. When a record contains several of these localizing sequences, there is little evidence to indicate progressive changes in their length according to their serial order. This is somewhat contrary to expectation. One would anticipate that, after a couple of localizing attempts, the technician already would have taken many of the readings necessary for narrowing the trouble to one stage, and that localizing sequences would reduce in length as the problem proceeded.

Prediagnostic Behavior

Before the discussion of component replacements, let us consider the amount of checking activity which precedes the initial replacement in each problem. There are good reasons why this portion of the record deserves separate analysis. For example, Damrin and Saupe (10), using the Tab Test, found the number of "prediagnostic"[40] checks was related to other problem characteristics. Also, our own data suggest that consideration of the activities occurring before the initial replacement are in accordance with the technician's conceptualization of his task.

For the average problem, about 14 prediagnostic checks are made. This number proves to be remarkably stable. It is

---

[40] i.e., The number of check tabs pulled prior to the pulling of a unit tab. The term "prediagnostic" will be used here in the interests of standardization, although it is not entirely apt since it refers to a segment of the performance which undoubtedly contains a major portion of the diagnostic behavior.

relatively unaffected by the type of equipment being repaired,[41] by the test format,[42] by the output symptoms,[43] or by the order in which the problems are administered.[44] Of course, there are wide differences between the number of prediagnostic checks for single performances.

What variables determine the length of prediagnostic sequences? Three factors were investigated. One of these was the difficulty level of the problems. Figure 3 on page 49 shows the average number of prediagnostic checks in each problem plotted against the proportion who solved the problem. As shown in the figure, there is a tendency for the average number of prediagnostic actions to be directly related to difficulty level. This trend holds for all but one of the problems. The "odd-ball" is problem C115. Without this problem, the rank correlation between the two variables is .73; with C115 included, the correlation drops to .39. Problem C115 has several unique characteristics which make this drop readily interpretable. Foremost among

---

[41] The average number of prediagnostic checks for the radio is 13.3; for the radar, 13.6.

[42] The average number of prediagnostic checks for Job Sample problems is 16.5; for the MASTS problems, 14.7; and for the AUTOMASTS problems, 12.9.

[43] The set of problems employed in this study have very few repetitions of the same initial output symptoms. As a result, analysis based on symptom types are fragmentary. However, on the basis of information available there is as much variability (in terms of prediagnostic checks) among problems with like symptoms, as among problems with unlike symptoms.

[44] Chi square tests of significance for the Job Sample data and the MASTS data indicated that there was no significant relationship between the order of problem administration and the number of prediagnostic checks. For details of this analysis, see Table E of Appendix C.

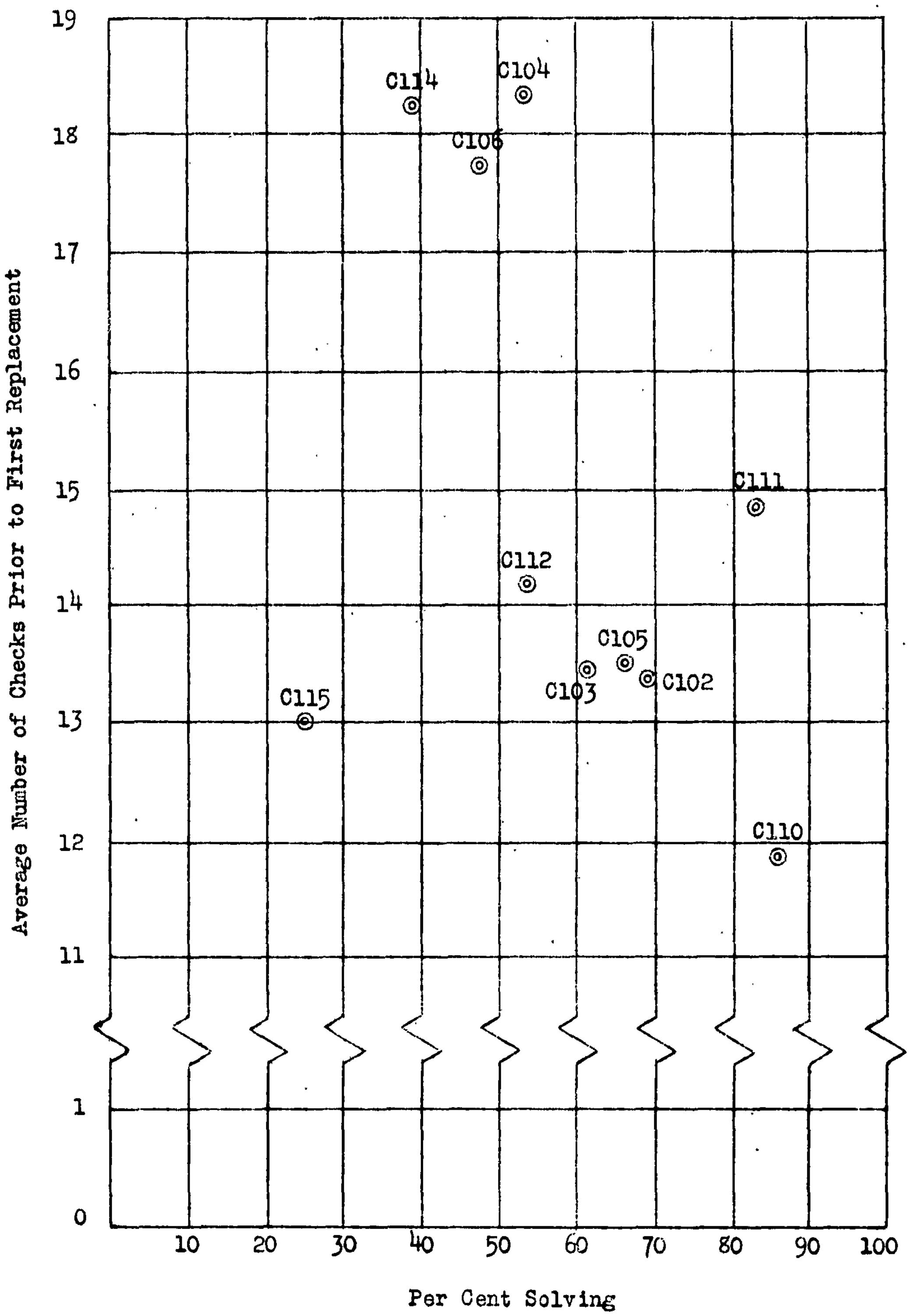


Fig. 3. Relationship Between Amount of Prediagnostic Activity and Difficulty Level.

these is the fact that it has clear-cut localizing cues which lead to the wrong stage of the equipment. As a result, most technicians quickly (and erroneously) isolate the difficulty to the step counter and replace the step counter tube. Barring this single exception, the overall relationship indicates that many difficult problems are difficult from their outset (i.e., before the first replacement). It has been suggested that these early difficulties may be due to the subtle nature of the cues (e.g., near-normal meter readings) available to the technicians. Trouble shooting researchers ought to be concerned more about determining exactly how various factors contribute to making a problem hard or easy. At present, it is impossible to specify with any degree of rigor the effects upon trouble shooting performances of such variables as complexity, ambiguity, type of cue, the extent to which cues vary from normal, the number of cues available, etc. Such variables deserve more systematic attention than they have received.

The second question investigated was the relationship between the length of the prediagnostic sequence and the type of initial attack sequence (IAS) adopted for the performance. It will be recalled that the number of actions in the total performances and the number of actions in the ILS's were related to IAS type. The result was that the IAS ordering according to length of total performance was different from the IAS ordering for ILS length. A parallel analysis between the length of prediagnostic sequence and IAS type resulted in a still different

ordering. Here again, we find that as a group, subjects with the tendency to be efficient require fewer actions to solve a problem, but each man "saves" actions in his own way. It is not possible to find a standard phase of the performance where the "efficient" subjects accomplish their savings.

The third relationship explored was between the number of prediagnostic actions and the order of the problems in the series. It was hypothesized that technicians would become more wary as the problem series proceeded, and that they would demonstrate this wariness by progressively increasing their number of prediagnostic actions. Chi square analyses failed to support the hypothesis; the amount of prediagnostic activity on the sixth problem was about the same as on the first.[45]

Up to this point we have discussed the possible determiners of the length of the prediagnostic series. A related question is whether dependable predictions can be made from the length of the prediagnostic series. In other words, "If you know the length of the prediagnostic series, what else can you say about the performance?" Damrin and Saupe (10), using their Tab Tests, found a small, significant correlation between the total number of prediagnostic checks and the total number of problems which were correctly solved by the first "unit pull." A parallel analysis was conducted using 700 performances from all formats. The

---

[45]Job Sample and MASTS data were employed in the investigation of this question since complete counterbalancing of problems and problem-orders were features of that design.

correlation coefficient of .05 indicates that there is no significant relationship between the two variables for these data.[46] Generally speaking, there is no advantage to a large number of prediagnostic checks *per se*.

One might expect that a long prediagnostic series would be associated with a relatively large number of actions following the initial replacement. The correlation (-.13) indicates that there is no pronounced tendency for the length of the two portions of the trouble shooting performances to vary together. There is, however, a definite tendency for an ET who has a relatively large number of actions before his first replacement on one problem to have a large number of pre-first-replacement checks on his subsequent problems. Similarly, the technician who has few prediagnostic checking actions in one performance, tends to have few in his other performances. This inter-problem stability is indicated by an Ebel coefficient of .70.

During the typical prediagnostic series, 1.7 isolating sequences (IS's) take place. In other words, the average technician does not substitute a component immediately following his first excursion into a stage. Instead, after his first specific intrastage check series, he either resumes his generalized localizing behavior or moves on to another stage. In 55 per cent of the performances the ET returned to the stage in which the original

[46] Negligible relationships also were obtained when the prediagnostic count was limited to checks in the stage where the first component was located.

isolating behavior occurred in order to make his first component replacement. About a third of his actions preceding the first replacement are in the same stage as that replacement.

The Replacement of Components

The replacement of a component serves as a "payoff." It usually represents the culmination of a particular line of attack. Stated in problem solving terms, a component replacement is the "verification" (47) phase wherein the technician tests the validity of the conclusions he has drawn from the preceding complex of symptoms and checks. In addition to being a critical test whose outcome is uncertain, each replacement represents a cost and therefore is something of a gamble. Aboard ship, the cost may be expressed in such terms as equipment downtime, technicians' time and effort, actual dollar value of expended parts,[47] and the reduced capability of the ship while the equipment is out of order. In the MASTS and AUTOMASTS testing situations, each replacement costs the technician 30 seconds of testing time. The significance of this type of action, both in terms of its logical importance and its cost, indicates that it deserves intensive treatment.

---

[47] A considerable proportion of the parts removed on the job are not defective at all (6). Furthermore, in actual shipboard practice good components are removed and often not put back into service.

Initial replacements are discussed prior to the treatment of replacements in general. There are two main reasons for this. One, almost every performance has an initial replacement. Two, the initial replacement is of special interest because it is the first critical test of the technician's skill in interpreting the information complex associated with a particular problem.

There are five types of components that can be replaced in each equipment. Figure 4 gives the popularity of each component type.

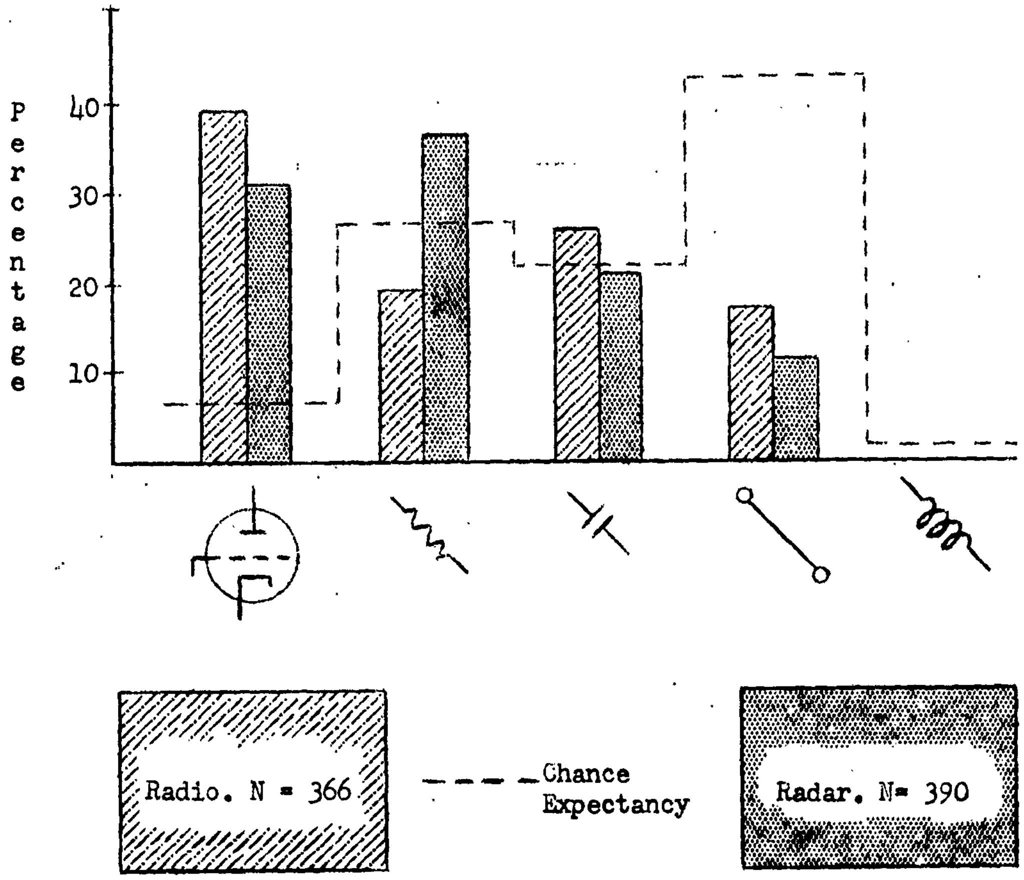


Fig. 4. Popularity of Component Types for Initial Replacement.

An examination of the results shown in Figure 4 on page 54 reveals that vacuum tubes are the most popular radio first replacements. Capacitors, resistors, and connectors are represented in that order. No one replaced a coil or a transformer at the time he elected to make his first substitution.

The radar data features the same general pattern, although there is a far greater tendency for initial replacement of resistors.[48] It is likely that this apparent inter-equipment difference is the result of specific problem conditions rather than a truly differential tendency induced by the type of equipment itself.

The dotted line on the figure shows the chance expectancy for the type of initial replacement. This is given as a single line (rather than one for the radio and another for the radar) because the proportion of replaceable components of each type was held approximately constant for each gear. It will be noted that the observed replacing behavior differed rather drastically from that which would have occurred had the initial selections been made blindly. For example, although only 7 per cent of the components which could be replaced were tubes, 38 per cent of all of the initial radio replacements fell into the

---

[48] One format difference deserves mentioning here. The ETs working with actual radar equipment prefer to replace tubes before trying any other type of replacement. Those who worked on the symbolic radar generally prefer to make a resistor replacement first. This may be due to the relative ease of tube substitution in the Job Sample test.

tube category. Also, although 43 per cent of the replaceable components were in the "connector" category, fewer than 20 per cent of the initial replacements involved members of that class.

One can only speculate why, in relation to chance expectancy, tube replacements were so popular, and leads, jumpers, and fuses were so unpopular. Perhaps tube replacements can be attributed to the bad reputation of tube reliability. However, one would scarcely advance the same reason to account for the relative paucity of connector replacements.

Figure 5 on page 57 shows the location of the radio initial replacements in terms of the functional layout of the equipment. The stages are arranged on the chart from left to right according to their order on the signal path. It is clear that most initial replacements occur near the middle of the set. The detector circuit, which serves as a link between the RF - IF section and the audio stages, is the most popular. Technicians were almost always able to eliminate either the front (antenna) or the back (speaker) end before they made their initial component substitutions.

The radar picture is quite different. As illustrated in Figure 6 on page 58, the usual first replacement occurred at either end of the set rather than in the middle. The concentration of initial replacement activity in the multivibrator, sawtooth generator, and oscilloscope stages is explained by the presence of actual defects in those stages. The popularity of the diode step counter presents a different interpretative problem, however, since there were no defective components in that stage.

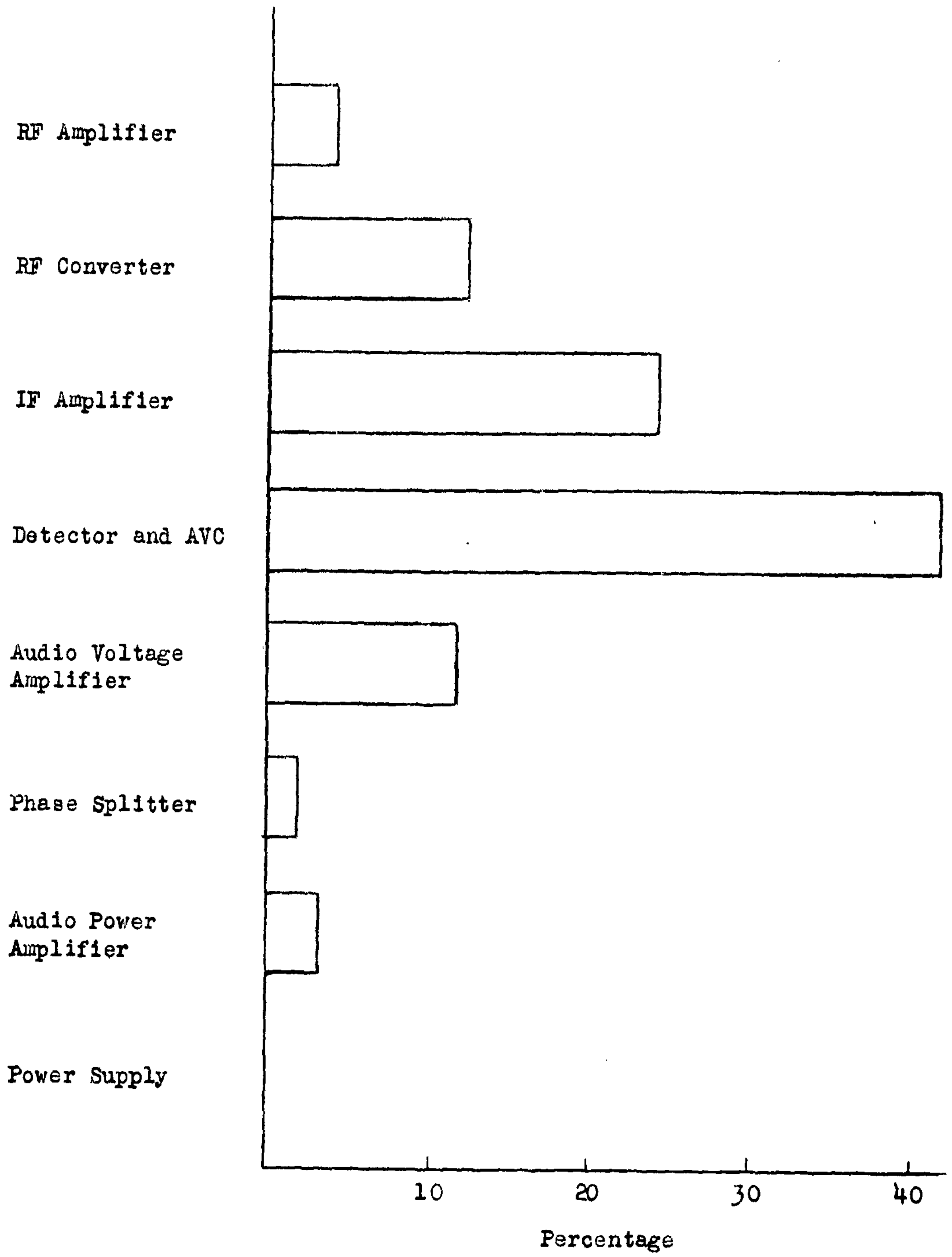


Fig. 5. Location of Initial Replacements in Radio Receiver. This figure is derived from Appendix Table G. N = 366.

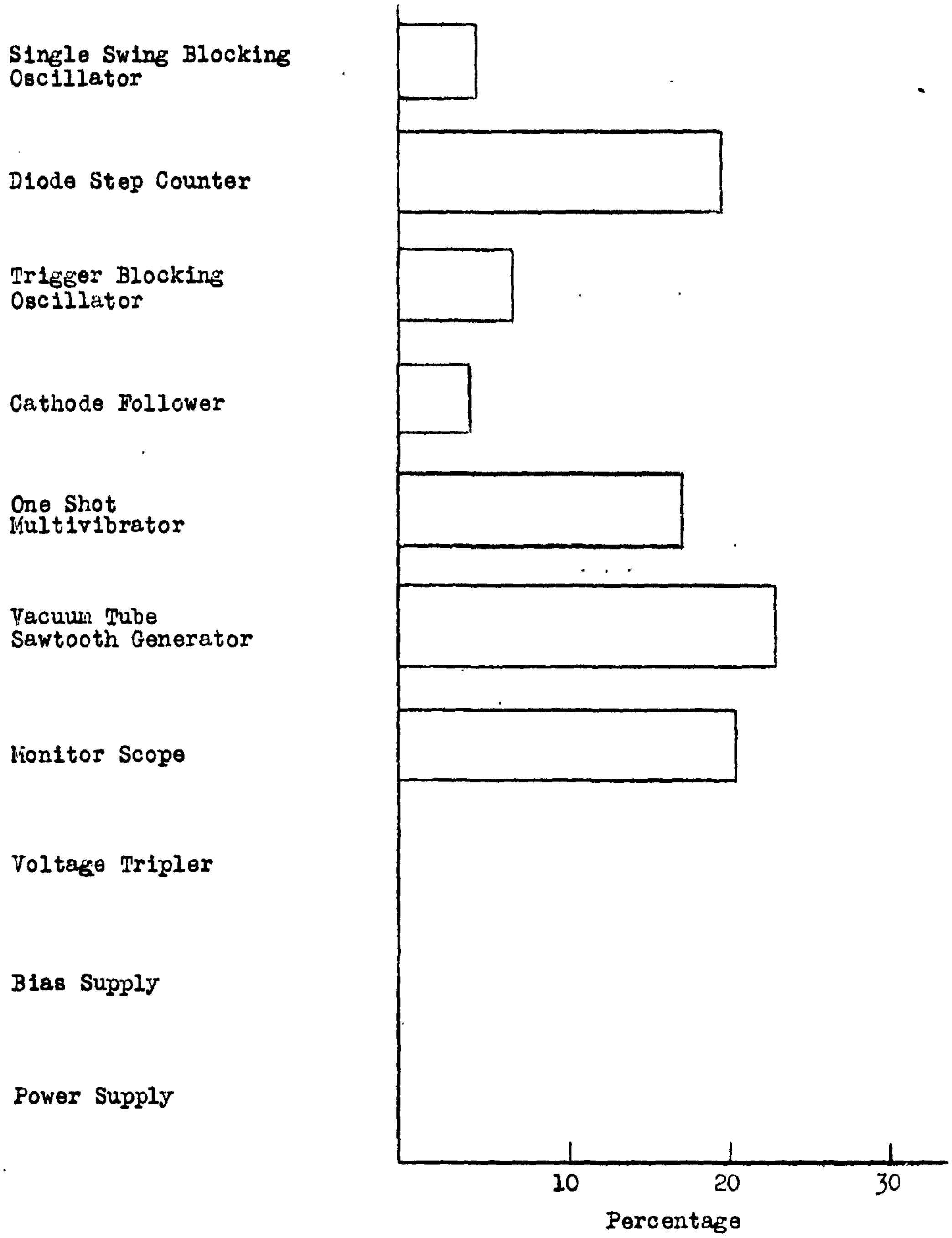
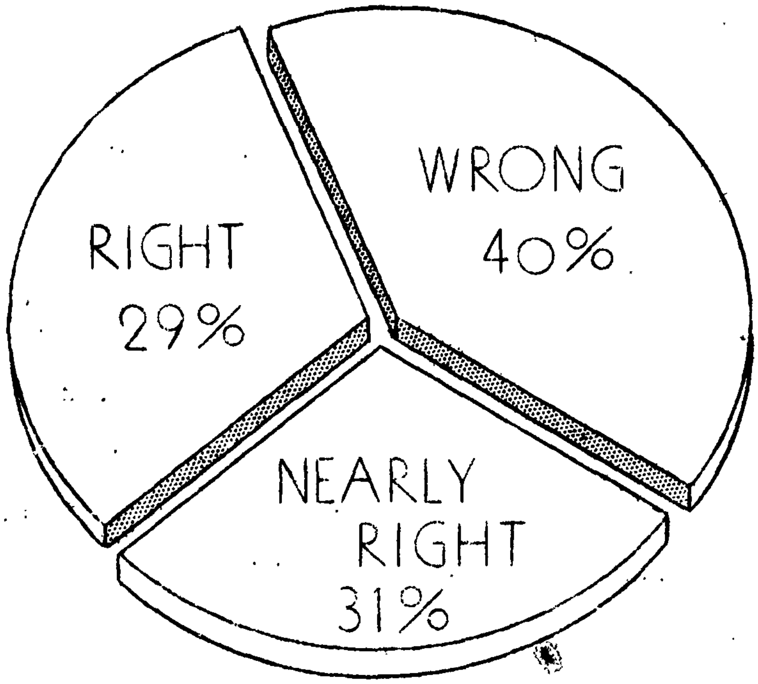


Fig. 6. Location of Initial Replacements in Radar. This figure is derived from Appendix Table H. N = 390.

Practically all technicians failed to appreciate that the step counter is "sensitive" to certain malfunctions in the triggered blocking oscillator, which follows the step counter along the pulse path. For example, if the bleeder resistance for the cathode bias of the blocking oscillator is changed, this "reflects back" into the step counter and prevents the step counter capacitor from discharging. If the technician trouble shoots by taking waveforms on a test oscilloscope, he obtains a "good" waveform on one side of the step counter tube and a "bad" (no vertical steps) waveform on the other side of the tube. Ordinarily the technician assumes that the trouble lies in the step counter, and he does not consider the possibility that a defect in the later stage could give a superficially clear-cut localization to the earlier stage. The best men eventually overcome the formulation of the problem in these strictly stage-by-stage terms and investigate other stages for additional clues. The assumption of relative functional independence between stages is generally valid in radio trouble shooting, but does not apply in the more complex inter-actions of pulse forming circuits. This seems to be a straightforward instance wherein technicians fail to recognize the possible limitations of "block" style trouble shooting.

How successful were the first replacements? How many of them restored the equipment to normal operation? How many of the initial component replacements were wrong, but close? Let us look at Figure 7 on page 60.



N = 769 Performances.

Fig. 7. Effectiveness of Initial Replacements.

As the figure indicates, 29 per cent of the initial replacements result in correction of the difficulty (i.e., the problem is solved on that replacement). Thirty-one per cent of the initial replacements are "near-misses"; while they did not eliminate the cause of the malfunction, they involved some component in the same stage as the faulty part. The remaining 40 per cent were outside of the stage in which the trouble was located. Are the "near-missers" more likely to be successful in the long run? The answer is affirmative; over half of the

performances (61%) which feature an initial "near-miss" terminate in eventual success. In contrast, less than a third (30%) of those with an initial wrong-stage replacement solve the problem.

There is another way of examining the effectiveness of the first component replacement; this involves the use of a <u>basic parts rating</u> system. Parts ratings were developed in connection with the "Component Replacement Score" which was reported previously (5). Each part is classified as either (a) the faulty component itself, (b) a component which is plausible on the basis of the symptoms and check information provided, or (c) a component which is not a reasonable replacement under the conditions of the problem.[49] Frequency analysis of each basic parts rating class revealed that 33 per cent of the first replacements were plausible and 38 per cent were not plausible. These results, though derived on a different system of evaluating effectiveness, are very similar to the "near-miss" frequencies shown in Figure 7.

The basic parts rating data also were used to investigate the relationship between the plausibility of the first replacement and the difficulty level of the problem. On the harder problems, one would expect fewer first replacements to be either correct or

---

[49] The ratings were established by expert judgments. For each problem, the judges were given all the information potentially available to the subjects, and were required to rate each replaceable part. Their results indicated that there were plausible replacements outside of the defective stage and implausible ones within the stage.

plausible. To investigate this, a correlation was computed between the average parts rating score on a problem and its difficulty level.[50] This treatment yielded a coefficient of .85. The data were treated further to determine the extent to which this high correlation was due to the "correct" component class alone. When the initial successes are eliminated from the average basic parts ratings, there is no correlation between the proportion of plausible replacements and the difficulty level of the problem. It can be stated, then, that the proportion of subjects solving the problems on their first replacement is highly related to the proportion of the group eventually solving it, while the proportion of plausible and implausible first replacements is not predictive of difficulty level.

In the preceding, the first replacements were singled out for special treatment. Over two-thirds of the performances had additional component substitutions. The following is devoted to an analysis of all substitutions.

Although a few technicians substituted as many as 17 components in the course of a single problem, such extremes were unusual. The average repair attempt included between two and four replacements, or about one replacement for every 12 checking actions.[51]

---

[50] A mean parts rating score was obtained (by problem) for all first replacements, using weights of 1, 2, and 3. This was correlated against the percentage of the group solving each problem.

[51] The number of replacements varied somewhat according to type of equipment and format. The Job Sample had more replacements and more checks per replacement. Also, there usually were more replacements in a radar problem than in a radio problem.

One might expect that some men would be prone to make substitutions quite freely, and that others would be chary of premature replacements. "Sea stories" and the good-natured banter of technicians would support the notion that ETs could be classified as either consistent "replacers" or consistent "non-replacers." If the technicians were indeed consistent, the replacers would be expected to accomplish a relatively large number of substitutions on each problem while the non-replacers would be frugal in changing components. A procedure devised to investigate inter-problem reliability indicates that no significant consistency exists.[52]

When the sample of 2259 parts is considered one finds that initial replacements are rather typical of replacements in general. (See Table F, Appendix C). There are two exceptions; one in the radar problems, and one in the radio series. The high popularity of initial tube substitutions during radio problems is not reflected in the count for the whole performance; the 38 per cent original popularity drops to 24 per cent when all pulls are considered. On the radar gear, the principal trend seems to be a reduction of inter-type popularity difference. The overall radar

[52] In order to partial out the effects of success on number of replacements, the performances were split into successful and unsuccessful groups, by problem. The median number of replacements was determined for each sub-group. In each problem sub-group, those performances above the median were assigned a plus and those below were given a minus. Thus, a perfectly consistent technician's score for the whole series would consist of either ten pluses, or ten minuses. A frequency distribution was made of the number of pluses assigned to each man. The obtained distribution of pluses was compared with a theoretical chance distribution by means of a chi square test.

replacement data indicate that tubes, capacitors, and resistors are about equally popular, with connectors finishing a strong fourth. The men continue to show very little inclination to remove coils or transformers.

A breakdown of replacements according to the various equipment stages also was made. This tabulation shows that there are no major differences in relative popularity of the stages between the total tabulation and that reported on pages 57 and 58 for the initial replacements. In the radio there is still a pronounced preference for components toward the middle of the signal path, while the ends of the radar pulse path continue to be favored.[53]

We saw that 60 per cent of the initial substitutions were in the same equipment stage as the defective component. It was hypothesized that the proportion of the replacements in the same stage as the defective component would vary inversely with the ordinal value of the replacements (i.e., the proportion of <u>first</u> replacements in the defective stage would be greater than the proportion of <u>second</u> replacements in that stage, etc.). Generally speaking, this hypothesis was verified with respect to the first, second, third, fourth, and fifth replacements. However, there were some exceptions,[54] and the trend was not as clear-cut as

---

[53]A complete tabulation of the replacements upon which these analyses were based is presented in Appendix Tables G and H.

[54]For example, on AUTOMASTS radio problems, there was a definite increase in the proportion of replacements in the defective stage as a function of the order of the replacements.

expected. When a parallel analysis was conducted on the basic parts ratings of successive replacements, essentially the same finding was obtained. These results indicated that the technicians usually had a general idea of the location of the trouble, as indicated by their tendency to remain in the neighborhood of the defective part. There was very little support for the notion that replacing behaviors became increasingly wild and unreasoned as the performances proceeded.

Despite this tendency to continue to substitute parts near the defective component, the likelihood of success progressively diminished with additional replacements. Figure 8 on page 66 presents three different bases used to estimate the likelihood of success.

Figure 8 consists of three graphs. The one labelled "A" shows that a decreasing proportion of the total group solved the problems at each successive replacement. The curve indicates that the likelihood of correcting the malfunction after one had replaced two or three components (unsuccessfully, of course) was slight. Since 29 per cent of the performances terminate with the first replacement, the form of the tail of the curve becomes increasingly dependent upon the changing composition of the subject group. In order to eliminate this factor, graph B was drawn using only those performances which actually terminated in success. Therefore, the percentages shown on the "B" graph represent the proportion of the successful performances terminating at each ordinal replacement position. This treatment again shows clearly

A
All Performances

B
Successful Performances
Only

C
Performances Having
an n[th] Replacement

Fig. 3. Success as a Function of Ordinal Position of Replacement

the diminishing likelihood of success with continued component replacing. Over half of those performances which end with the correct component replacement, have only one replacement. If a technician has not been successful in his first three replacements, there is little reason to be optimistic of his eventual success. To rule out the effects of the variable number of replacments at each ordinal replacement position, the proportion of successful replacements was computed for each replacement postion separately (i.e., "How many of the performances which had an nth replacement, had a correct nth replacement?"). As shown in graph C, the chance of success decreases sharply as the ordinal position of the replacement increases. These findings suggest the possibility of a "three strikes and you're out" practical limit for problems of this nature. Instead of rigid time limits, it may be realistic to allow a man to continue working either until he has used up all of the time allotted for that particular problem or until he has accomplished three component substitutions.

Sometimes a man would make several replacements before attempting to obtain further test equipment information regarding the circuit.[55] An inspection of the records in which three or more consecutive replacements were made indicated that these substitutions

---

[55] The average length of an uninterrupted replacing sequence was 1.3 components; sequences of five or more occurred with less than one per cent frequency.

were almost always clustered around a single point in the equipment. If they weren't all in a single stage, they were apt to be centered around the coupling between adjacent stages. While several interpretations may be offered to account for consecutive replacements,[56] such sequences unquestionably are <u>not</u> tests of separate and distinct trouble shooting hypotheses.

On rare occasions, a record would show a series of consecutive tube substitutions that suggested that the ET was systematically working his way through all the tubes in the set. It had been anticipated that this type of behavior might be very common, perhaps a usual prelude to more analytic attempts to locate the difficulty. Without attempting to speculate why they did not occur, suffice it to say that replacing sequences of this type were too infrequent in the present data to warrant serious examination.

Although replacements within a consecutive series have a degree of mutual relationship, such is not the case with respect to replacements which occur singly in the performance records. Attempts to predict the type of component that will be replaced next in a given record, on the basis of knowledge of the types of components

---

[56] Perhaps this type of behavior simply represents the technicians' conviction that the defective unit is in the restricted area from which the components are being drawn. This view is supported by the fact that such clusters of pulls frequently occur after an earlier replacement in the same area of the equipment. According to this interpretation, the ET localizes the trouble to a specific area (maybe erroneously) and makes a replacement calculated to correct the trouble. When this fails, he investigates other possible areas and decides that his earlier localization was right. Hence, he returns with renewed conviction and substitutes a number of the components in the suspect stage (or area).

already replaced during the performance, are not very fruitful.[57] There is practically no evidence of patterning among those replacements which are interspersed in the records nor do the men have consistent individual preferences for a given type of part. On the basis of this information it seems that replacements are made with reference to preceding checking actions, rather than previous replacing actions.

The Pre-replacement Block

The three actions preceding each component replacement were designated as the pre-replacement block and were singled out for special analysis. The general purpose was to determine relationships between the activities immediately before a replacement and characteristics of the replacement itself. One basic issue was to find out if it was possible to anticipate the removal of a component. Was there something about the way that the technician was behaving which signalled that he was about to make a substitution?

One characteristic of behavior occurring within the three-action sequence prior to a parts replacement is outstanding--its tremendous diversity. There is no standard pre-replacement procedure. Only one general statement seems to be supported by the

---

[57] It is somewhat more likely that a given type of component will be followed by another component of the same type than by one of a different type. For example, tubes are more likely to be followed by tubes than by any one of the other types of components.

analysis. The technicians almost always precede a component replacement with a series of checking actions using a single type of testing equipment. To explain more fully, they make a number of signal injections, oscilloscope readings, voltage readings, or resistance checks before the substitution of a part, but they rarely intermingle these different types of checks just prior to a replacement. This type of activity seems sensible enough when one realizes that the usual strategy of the trouble shooter is to make a substitution only after test equipment readings on each side of the component appear to be out of harmony with each other. On the basis of this information the only general prediction about component replacing behavior that can be made from knowledge of the pre-replacement behavior is: if a technician switches from the use of one type of test equipment to another, it is unlikely that he will replace a component within the three actions.[58] Attempts to predict the type of replacement made, its success, and other characteristics of the performance records from knowledge of the actions taken during the pre-replacement block were unsuccessful also. While the diversity of behavior in that block does not readily lend itself to dependency analysis, it is often quite sensible when reviewed in the light of all of the performance which precedes it. This implies that a larger

---

[58]In the case of certain specific patterns, positive predictions can be made. For example, if a man makes three or more resistance readings in a row, it is likely that he will make a replacement within the next three actions.

segment of performance is required in order to "understand" the component-replacing actions. It is another indication that the type of trouble shooting examined here is not a simple stochastic process.

The Post-replacement Block

Ofttimes a component is substituted but symptoms of malfunction persist. In these cases, the technician presumably has employed the best means at his disposal, and yet he has not solved the problem. Instead, he is faced with a new, more difficult one: i.e., the obvious suspect is not defective--a new possibility must be found. How does he regard the situation at this point? What does he do?

On 30 per cent of these occasions the ETs continue to check the stage where they made the component substitution. For the most part, their continued checking involves DC voltage readings and resistance readings. Since the average incorrect replacement was, in fact, near the defective part, this course of action was "statistically" prudent.

Twenty-one per cent of the post-replacement blocks[59] were devoted to confirmatory behavior. In these cases, the technicians verified checks that had directed them to the stage in the first place. They seemed unwilling to give up the hypothesis that the

---

[59]The post-replacement block generally consisted of three to six actions. If several replacements were made consecutively, the block began at the end of the last replacement in that series.

trouble was in that particular stage and appeared to be disinclined to seek elsewhere.

Twenty-eight per cent of the post-replacement blocks consisted of voltage and resistance checks in another stage. Frequently, a replacement in another stage was made within a very few actions. This behavior was construed to mean that the ET had entertained an alternative hypothesis at the outset; i.e., he had isolated to <u>two</u> stages rather than one.

In 16 per cent of the cases, an unsuccessful component substitution was followed by behavior indicating that the men were looking for another stage to which to localize. Typically, they engaged in signal-injecting and signal-tracing techniques of the same general nature as those identified with the beginning of a performance. One gains the impression that the stage in which the initial replacement was made has been abandoned (at least for the time being).

The remaining instances (5% of the total) were not easily classified. They did not fit well into any of the above four types of post-replacement activity, and they were not entirely homogeneous. Most of them tended to be rather erratic, i.e., they jumped from stage to stage in an unpredictable manner.

The technicians varied the type of activity engaged in following a component substitution; indices of consistency within a problem or across problems were very low. Apparently the post-replacement reaction is a "spur of the moment" matter, and it is not clearly related to personal styles of trouble shooting.

Brief Review of the Trouble Shooting Schema

Sections of this report may have been quite tedious. This is an unfortunate consequence of dealing with seven sets of data, obtained from four sources. It also is caused by the authors' desire to specify completely the analytic procedures applied (so that the reader will not be led to overgeneralize the results obtained). In spite of these necessary precautions, it is not all trees; there is a forest. It is appropriate at this time to review the trouble shooting schema, filling in the best single numbers available[60] in order to obtain an overview of the typical trouble shooting effort.

Our average electronics technician begins his repair efforts by pondering the gross output of the equipment and studying the schematic diagram. After a bit of this, he fiddles with the front panel control knobs.[61] Reassuring himself that these are properly adjusted, he hooks up a signal generator[62] or an oscilloscope.[63] With these he injects signals or takes waveforms along the principal data flow chains.[64] His checking is rather systematic and proceeds from the rear end of the set toward its front. He intersperses a

---

[60] In this summary, the values reported represent the average obtained at each point in the trouble shooting sequence where information was available.

[61] On the radio, he manipulates the volume controls; on the radar, the centering controls.

[62] When working on the radio.

[63] For use with the radar equipment.

[64] The main signal path on the radio; the pulse path on the radar.

couple of B+ readings at the power supply to make sure that he is getting suitable power. Early in a radar performance (within the first seven or eight actions) he checks the output of the master oscillator. On the radio, he quickly dispenses with the audio section. His real decisions lie ahead.

Along about his eighth check he finds (or thinks he finds) a break in the data chain. This leads him to a specific stage[65] of the equipment. At this point a phase shift occurs in his trouble shooting activity. He disconnects the signal generator (or scope) and rigs up a multimeter. Then he checks resistances and DC voltages within the suspect stage.

He makes three or four such checks. He leaves the stage without replacing a component, although there is a better than even chance that the defective part is in that stage. The signal generator (or the scope) again is brought into play. He checks the signal at several places near the point where he originally noticed the signal break. After three or four confirming actions of this type he returns to the stage.

Without further checks, he pulls the tube and replaces it with one from spares. The symptoms persist. He takes two or three resistance readings in the stage and then replaces another component. This time he succeeds in correcting the trouble.

[65] On the radio, the detector; on the radar, the vacuum tube sawtooth generator.

SECTION IV. SPECIFIC TROUBLE SHOOTING ISSUES

There are many specific issues regarding trouble shooting behavior which deserve special treatment. Six of them are examined in this section. The first investigates how well technicians apply the check information that they accumulate. Involved here is the definition of a theoretically "sufficient" combination of checks for certain diagnostic purposes, and evaluation of the relationships between acquisition and utilization of these key checks. Matters of this type are dealt with under the heading "Utilization of Information."

The second issue concerns changes in "distance" from the trouble as a function of each successive action. This treatment includes a tryout of a new technique for assessing the rate of convergence to the locus of the difficulty.

The third issue is devoted to expert judgments of trouble shooting quality. The judgments previously had been used as an intermediate criterion for test scoring methods. In the present instance, the goal is to specify the correlates of performance on which the judges based their ratings. The approach adopted was to develop an objective successive sorting procedure which would reproduce the original expert judgments.

The concept of redundancy as it relates to trouble shooting is discussed next. Types of redundant behavior are differentiated and tabulated according to frequency and location within trouble shooting record. Also, the question of the positive and negative consequences of redundancy is discussed.

The next part considers the problem of errors in trouble shooting, and presents a description of their frequency of occurrence and overall significance.

The final issue covered here is a general discussion of the relationship between trouble shooting and the general psychology of problem solving. Emphasis is placed upon the contribution that each type of study can make to the other.

Utilization of Information

A basic premise of trouble shooting is that the technician or repair man can collect (by proper checks, tests, manipulations, etc.) enough information about the various operating characteristics of the equipment to allow him to identify and replace the malfunctioning part.

For every trouble which develops in electronic equipment, there are combinations of checks which are logically sufficient to indicate the stage or unit in which the trouble lies. Also, for every trouble, there is at least one piece of information (or combination of pieces of information) which will indicate, theoretically, which component is at fault.

All of this is apart from whether the technician finds these items of information and whether or not he recognizes them if he does find them. Knowing the exact trouble, it theoretically is possible to derive the various combinations of information which are sufficient to localize the trouble to the defective stage and

those that are sufficient to isolate the trouble to the faulty component. Then, it is a simple matter to follow the trouble shooter's steps in searching for the trouble and to determine the points in the performance (if any) where he has enough data to localize to the stage and where he has enough to correct the trouble. Such a treatment permits empirical determination of the extent to which the technicians succeed in finding and utilizing information which will aid them in solving the problems.

The following analysis was carried out on 422 AUTOMASTS performances, on ten radio and radar problems. The procedure was essentially as outlined above; that is, each performance was examined to determine if the trouble shooter had made a combination of checks from which he theoretically could deduce which stage contained the trouble. As soon as a sufficient combination of these checks was encountered in a given record, the record was marked at that point to indicate that the Optimum Point of Entry (OPE) had been attained. The record was further traced to ascertain whether sufficient information had been secured to give an absolute indication[66] of which component in the stage was faulty. A mark was placed at this point to indicate that the Optimum Point of Solution (OPS) had been reached.

[66]On several problems, the information available from the AUTOMASTS permitted isolation to one or the other of two parts. In these cases, attainment of this information was counted as sufficient.

Once the OPE and OPS are located--in those records which contain them--several tabulations give an empirically derived picture of how well the technicians utilize the information they obtain, and how well they succeed when they do <u>not</u> obtain the "required" information.

Three separate but related analyses describe performances in which the OPE and OPS are discovered and the results when those points are not found. The first is concerned with the average performance in which both the OPE and the OPS are reached. The second traces the patterns of performances for those who did, and for those who did not, encounter the OPE. The third analysis deals with the OPS in terms of occurrence and effect.

```
Start
  '
                          OPE                 OPS                 End
            7 actions      '      8 actions     '     7 actions      '
A                          '    1 replacement   '   2 replacements   B
```

Fig. 9. Location of OPE and OPS in Typical AUTOMASTS Performance.

Line AB represents the length of an average of the performance in which both the OPE and the OPS were present. From the start, it takes an average of seven actions to obtain sufficient information to make an optimum entry into the defective stage. From that point, it takes another eight checks (including at least one component replacement) to obtain sufficient information

to replace the faulty component. The average performance also contains an additional seven actions after reaching the OPS to complete the performance. The probability is high (.76) that the performance will be successful if both the OPE and the OPS are reached.

As is indicated by the seven additional actions (including approximately two component replacements) after the OPS is reached, the acquisition of this information does not lead to an immediate solution. Perhaps technicians feel a need for verification of the indicated solution (a third of the actions after the OPS are repeats), or more likely, they do not recognize immediately the implications of the information they obtain at the optimum point of solution.

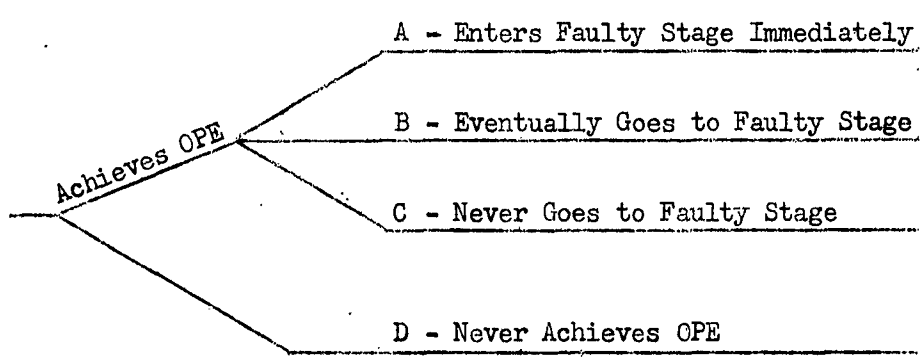


Fig. 10. Optimum Point of Entry Flow Chart: Identification of Alternative Routes.

The second analysis deals with 422 AUTOMASTS performances according to the flow chart shown in Figure 10. The explanation of the flow chart is straightforward. In some of the 422

performances, the technicians start checking in the faulty stage immediately after reaching the OPE. In the diagram these performances are labelled "A." Others make one or more checks after the OPE before entering the faulty stage (B). In still others, the technicians achieve the OPE but never perform checks in the faulty stage (C). Also, certain performances have no OPE at all (D). Each of these types will be discussed separately.

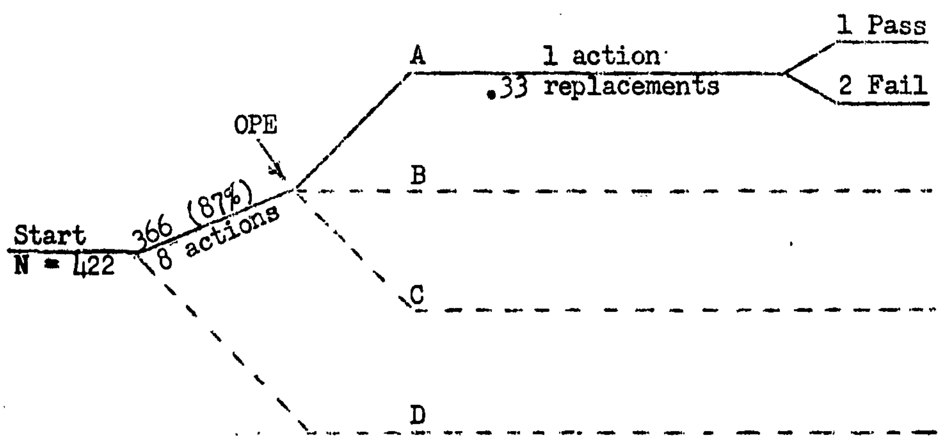


Fig. 11. Optimum Point of Entry Flow Chart:
Those Who Got to the OPE and Immediately Entered Faulty Stage.
Note: Throughout this discussion, the path being considered is traced by the solid black line.

In Figure 11, 87% of the original 422 performances reached the OPE. This required an average of eight actions. Of these 366 cases, only three entered the correct stage immediately. They had an average of one action after the OPE indicating that they terminated almost immediately. One of the three solved the problem and his was the only parts replacement after the

OPE. It can be seen that only rarely did a technician enter the stage where the fault was located immediately after the acquisition of sufficient information.

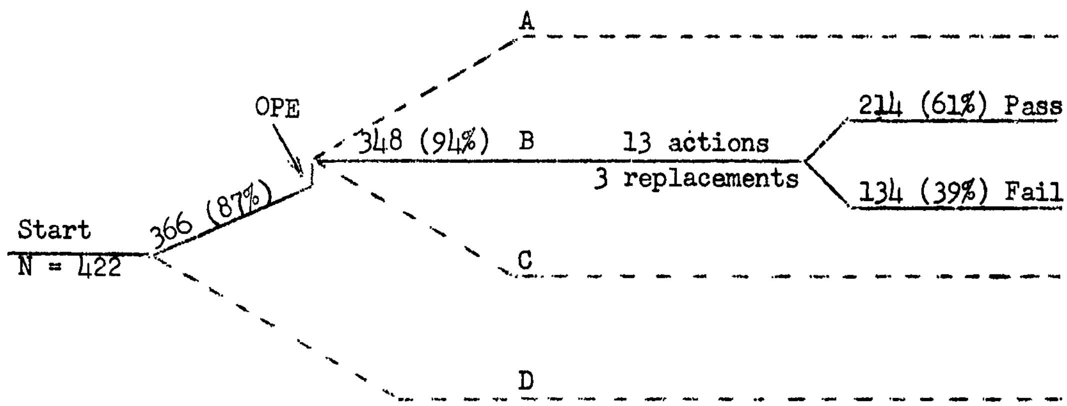


Fig. 12. Optimum Point of Entry Flow Chart:
Those Who Got to the OPE and Entered Faulty Stage Eventually.

Ninety-four per cent[67] of the performances which contained an OPE eventually (but not immediately) made checks in the faulty stage. From the time of reaching the OPE, typically this group had 13 more actions (of which three were component replacements) before the end of the performance. Sixty-one

[67]For these diagrams, each percentage is based upon the n before the preceding branching. For example, 94% is based upon 366 cases, not on the original 422 cases.

per cent of these performances terminated with the correct solution. This is approximately the same percentage as that found for the total AUTOMASTS sample.

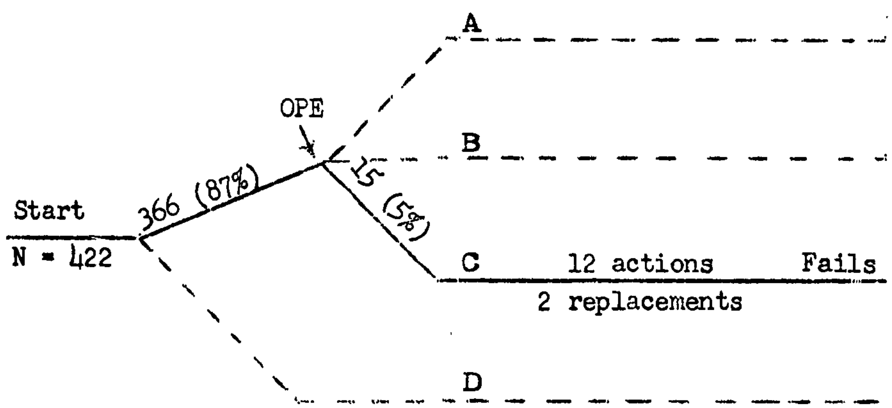


Fig. 13. Optimum Point of Entry Flow Chart:
Those Who Get to OPE But Never Go to Faulty Stage.

Only five per cent of the performance containing an OPE fail to have follow-up checks in the faulty stage. These performances (Type C) average 12 actions (two are component replacements) after the OPE to the end of the trouble shooting attempt. Of course, all of these ended in failure. Performances of this type are bad. Fortunately, they are relatively rare. Most ETs who obtain the information required to enter the malfunctioning stage, eventually do so. Type C attempts to signal a need for additional training.

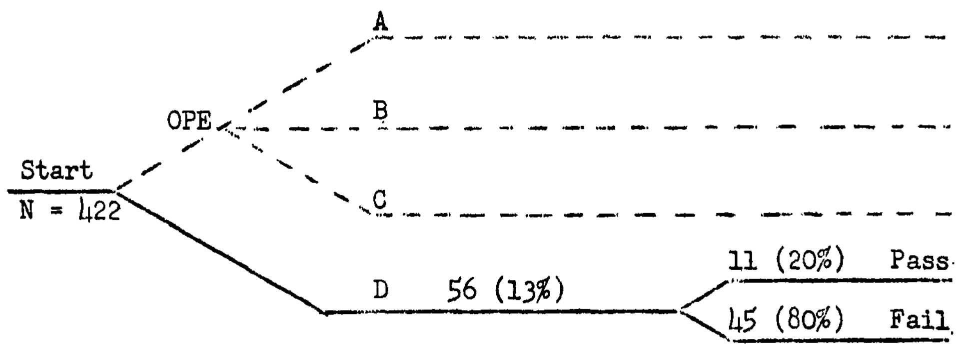


Fig. 14. Optimum Point of Entry:
Those Who Never Got to the OPE.

Of the original 422 performances, 13% do not contain an OPE, i.e., the ETs did not collect sufficient information to determine absolutely which stage contained the malfunction. This did not preclude the possibility that some of them might have entered the correct stage without sufficient information--many did. However, 80% of those who did not achieve the OPE failed.

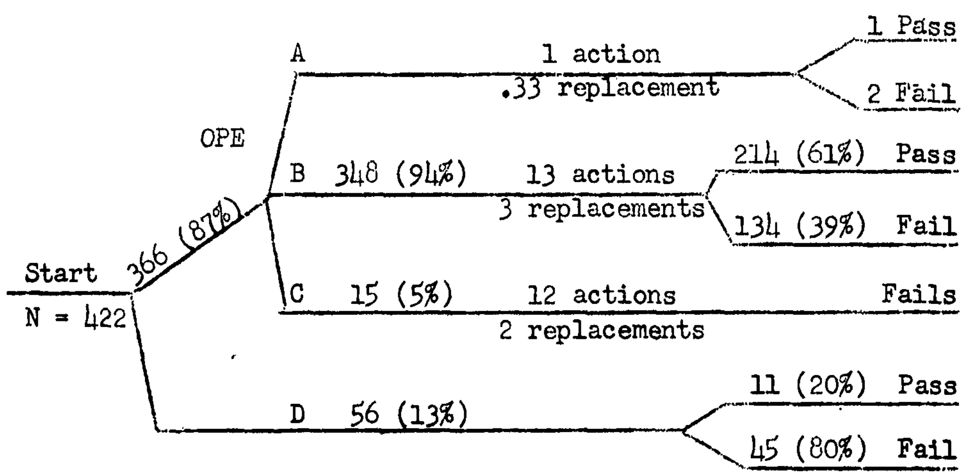


Fig. 15. Optimum Point of Entry:
Summary

Several points are brought out by the preceding analysis. The first is that a very large proportion of the 422 AUTOMASTS performances achieve enough information to locate absolutely the stage containing the malfunctioning part. Further, in most of these performances the technicians eventually do enter the correct stage.

Another interesting fact is that very rarely do those who achieve the OPE enter the stage immediately. It is highly probable that the information which theoretically is sufficient for immediate identification of the stage containing the trouble is not sufficient for the technicians tested on the AUTOMASTS. Although obtaining the OPE information does not insure success, approximately 60% of those who did get it were successful, while only 20% of those who did not obtain it were successful.

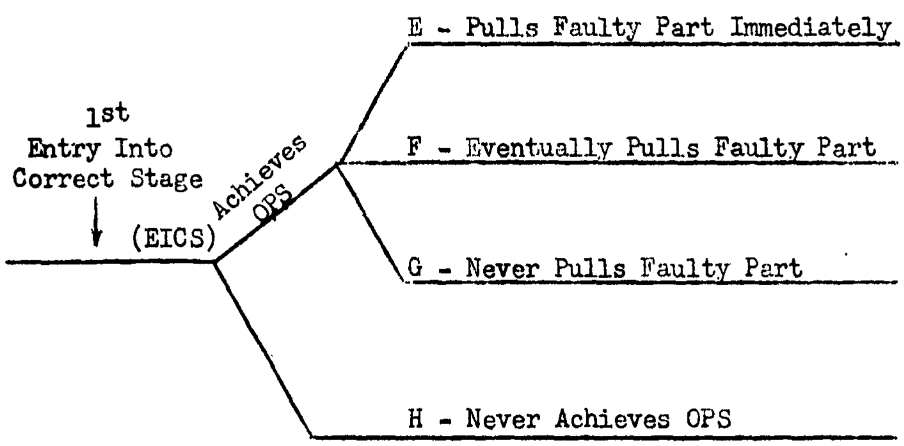


Fig. 16. Optimum Point of Solution: Identification of Alternative Routes.

The third type of analysis deals with the Optimum Point of Solution. The diagram in Figure 16 gives the four types of performances considered here. Interest is focused on the point of first entry into the faulty stage (EICS) and, therefore, considers only those performances which contain such an entry. The four types of performances include those where the malfunctioning stage was entered, the OPS was reached and the defective part immediately replaced (E); those in which the OPS was followed by one or more actions prior to replacement of the defective part (F); those in which the OPS was not followed by replacement of the defective part (G); and those who entered the faulty stage but did not reach the OPS (H).

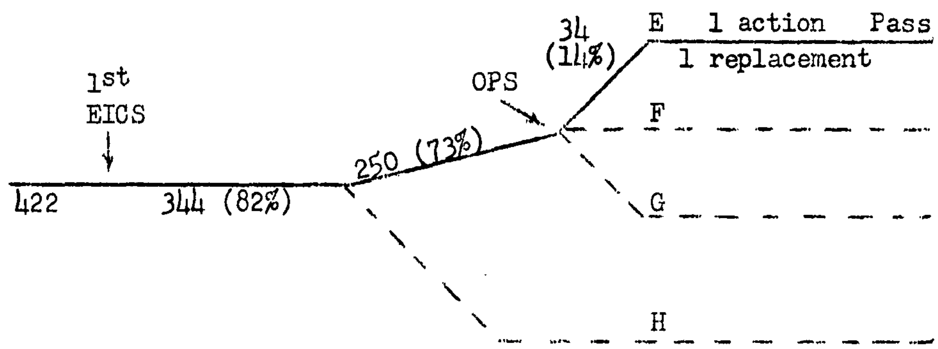


Fig. 17. Optimum Point of Solution:
Those Who Got to the OPS and Immediately Replaced
the Defective Component.

In this diagram, the solid black line indicates the performances in which the faulty part was replaced immediately after reaching the OPS. Of the original 422 performances,

82% entered the malfunctioning stage somewhere in the performance. Of these, 73% reached the OPS. Fourteen per cent of those reaching the OPS pulled the defective component immediately following the OPS.

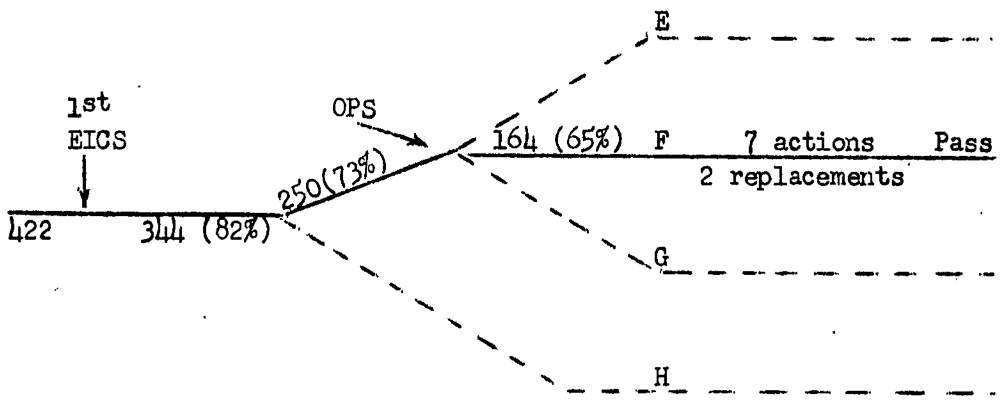


Fig. 18. Optimum Point of Solution:
Those Who Got to the OPS and Replaced the Defective
Component Eventually.

Of the 250 performances reaching the OPS, 65% eventually replaced the defective part. On the average, it took them seven actions to achieve solution. Here is substantiation of the fact that the theoretically sufficient information does not immediately lead to replacement of a faulty component. As a matter of fact, the average performance in this category contains several checks and two unsuccessful replacements following the action(s) which theoretically pin-point the difficulty but before the successful replacement.

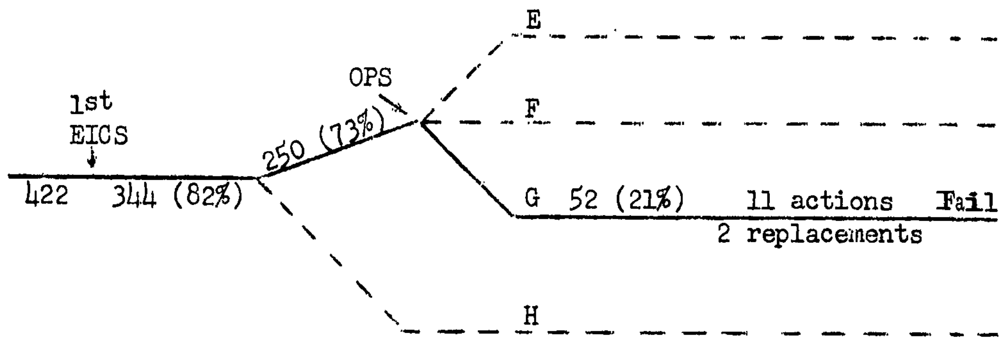


Fig. 19. Optimum Point of Solution: Those Who Got to the OPS But Never Replaced the Defective Component.

In the Type G performances shown above, 21% of those achieving the OPS did not recognize the clue and subsequently failed. The eleven moves from the OPS to the end of the performance is somewhat longer than for those who eventually solved the problems.

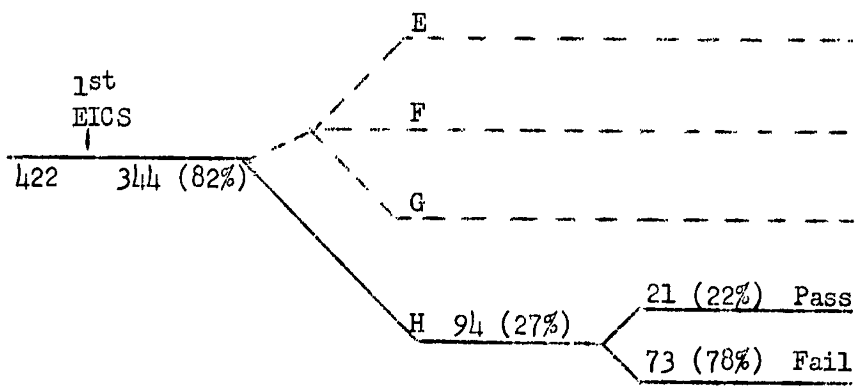


Fig. 20. Optimum Point of Solution: Those Who Did Not Get to the OPS.

Figure 20 on the preceding page shows those 94 performances in which an entry in the correct stage was made but a subsequent arrival at the OPS did not occur. Again, as in the similar case in the OPE analysis, lack of an OPS does not preclude a solution. About a fourth of those entering the correct stage did not attain the OPS, and 22% terminated successfully.

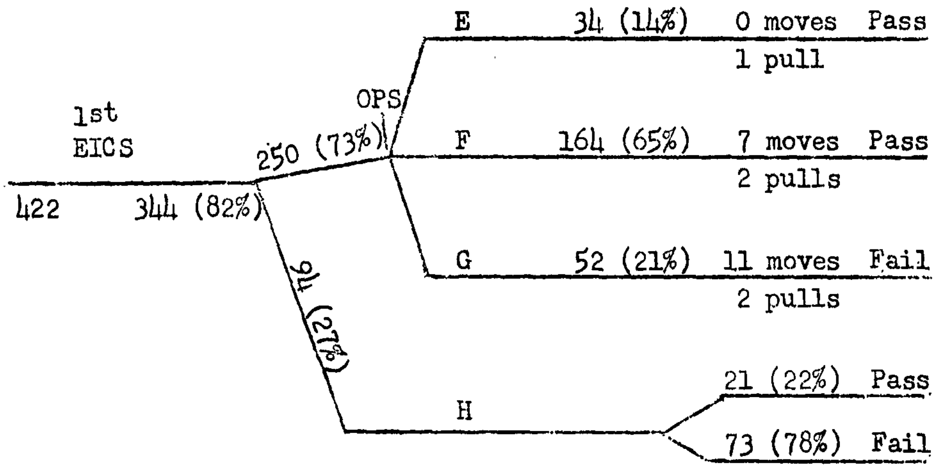


Fig. 21. Optimum Point of Solution: Summary.

To summarize the OPS analysis, several points stand out. For one thing, it is clear that the "sufficient" information is not necessary to the solution of the problems. Twenty-one of the 422 performances ended successfully even though they contained no OPS. On the other hand, it is obvious that gaining the OPS is a significant factor in success--over 90% of the

successful performances did reach the OPS point, whereas 58% of the unsuccessful performances did not achieve it.

After locating the OPS, only 14% of the performances made use of it immediately. However, a large percentage of the performances are solved despite this inefficient use of the check readings.

It is convenient to summarize here the results from the foregoing analyses:

1. A large proportion of the performances gain enough information to make, theoretically, an absolute identification of the malfunctioning stage and to identify positively the defective part.

2. In general, only a small proportion of those performances containing this information make use of it immediately.

3. Acquisition of the information necessary for the OPE and OPS does not guarantee a correct solution. However, it does appear to increase the chances of success.

4. About a third of the average performance is devoted to collecting sufficient information for the OPE, another third is given up to getting the OPS, and the remaining third apparently is spent in suplementing these theoretically sufficient clues.

It is important to note that technicians who worked on the AUTOMASTS problems, although some of the best of the Navy's trouble shooters, did not quickly recognize and act upon the information they derived from their trouble shooting efforts. Their acquisition of the information, based on these and other results given in this report, is not a serious problem. It is the utilization of the information that is the biggest deterrent to maximum efficiency,

and perhaps to maximum success. It may be that the training of future technicians should place greater emphasis upon this aspect of trouble shooting.

Rate of Convergence

According to one conception of trouble shooting the technician begins a problem at some distance away from the locus of the difficulty and progressively works toward it. Expressed in functional distance terms, he starts out at some more or less functionally remote part of the equipment and decreases the functional distance to 0 (if he solves the problem).

At this time let us briefly examine the data to determine the extent to which this tendency (to progressively reduce the functional distance) can be detected. If it occurs in trouble shooting records, let us explore its characteristics.

Fortunately, the Proximity Score described in Technical Report No. 12 of this series required the development of functional distance weights for each action of every AUTOMASTS problem. These weights run from 0 (for actions in the same functional unit as the defective part) to 5 for those which are functionally independent of the malfunctioning circuit. This system may be adapted readily to the present analysis.

The average initial action in an AUTOMASTS performance has a functional distance weight of 2.7. This indicates that most of the technicians begin checking at a point in the

equipment near the middle of the functional distance scale. From the first through the sixth action this average distance decreases regularly, at a rate of approximately 0.2 of a unit per action. By the time of the first isolation sequence the mean weight is down to 1.6. The average drops to 1.1 at the initial component replacement. This indicates a tendency for the ETs to focus more and more of their activity in the functional vicinity of the trouble as their performances progress.

Most (63%) of the changes from one functional distance zone to another are in the direction of the trouble (i.e., the "right" direction). Since different technicians begin a given problem at different points and all tend to end up in the neighborhood of the defective part, this type of activity is labeled <u>convergence</u>. It is then an interesting analytic problem to determine systematically the rate at which convergence occurs.

To do this each performance is plotted on a set of coordinates. The vertical axis represents the range of functional distance weights, and the horizontal axis represents successive changes in functional distance zones. The first point plotted is the functional distance weight of the initial action. From this point the record is traced action-by-action[68] and every

---

[68] On the average there are 2.3 actions for every change. The average length of "run" (number of actions between changes) is two. That the longest runs occur in the vicinity of the trouble is indicated by the fact that the average run in the <u>5</u> zone is 1.8 actions long, whereas the average run in the <u>1</u> zone is 3.3 actions long.

change (i.e., when the functional distance of an action differs from that of the action which preceded it) is plotted. Connecting the points gives a graphic representation of the convergence. Since in most performances, this plot appears to be linear, a straight line may be fitted by the least squares method (33). For each performance the slope of the best-fitting line indicates the rates at which convergence occurs. A steep slope indicates a direct convergence on the trouble, with few changes in the wrong direction. On the other hand, a performance which doesn't move consistently in the direction of the trouble has a negligible slope. Performances that consistently move away from the trouble have slopes in a negative direction. Thus, the Rate of Convergence (expressed as the slope of the line) provides a numerical way of expressing the convergent tendencies of each performance, and appears to have definite score potentialities.

To investigate the score potential, a sample of 100 performances was drawn from a pool of 422 AUTOMASTS performances. The drawing was accomplished by a stratified random procedure which insured that the sample represented the entire range of talent in the same proportions as the parent population. Each performance was scored according to the line-fitting procedures described above.

It is of particular interest to compare the Rate of Convergence score with the Proximity Score which is based on the same kind of a distance metric. The Proximity Score is an

average of the functional distance weights assigned to all actions in a performance. A high Proximity Score indicates that the average action within the performance is far away from the defective part. A low Proximity Score shows that the bulk of the performance is near the malfunctioning component. The Rate of Convergence introduces two new elements: a consideration of the changes in functional distance only, and a consideration of the particular order in which the changes are made. The Pearson correlation between Rate of Convergence and Proximity is low (-.18), indicating that the new score is relatively independent of the old, and that if it measures anything, it measures something different from the old score. A correlation between the Expert Judgments used as intermediate criteria in an earlier study (5) and the Rate of Convergence is significant at the .01 level ($r = .43$). Also, very significant relationships are found between the Rate of Convergence and the Direct Clue Actions score, the Number of Actions, and the Number Solved ($r = .44$, -.56, and -.47, respectively). These correlations suggest that Rate of Convergence has several characteristics to recommend its further consideration as a score for trouble shooting performances.

Generally speaking, this analysis substantiates the convergent conception of trouble shooting. The technique developed to explore convergence appears to have definite scoring possibilities. Also, the present analysis demonstrates that conventional averaging procedures may conceal informative sources of variance contained in step-by-step trouble shooting performances.

A Successive Sorting Procedure for Estimating Quality of Performance

When the overall quality of a complex performance must be determined and objective criteria are not well established, it is often desirable to use grades assigned by expert raters. This course was adopted in earlier phases of the present research (5). It was found that experienced judges could, under appropriate conditions of experimental control, assign grades to detailed records of AUTOMASTS repair attempts. Remarkable unanimity among independent judges was achieved, with coefficients of agreement above .90. These grades then provided a satisfactory criterion for evaluating objective scores. The only disadvantages of such judgments are the administrative considerations involved in keeping highly trained judges at a demanding task for long periods of time.

A number of scores were highly correlated with the judgment grades; at least six measures had coefficients in the .60 to .85 range. Relationships of this magnitude are satisfactory for prediction purposes, but they probably do not reflect the procedure on which the judges accomplished the grading task. For example, the Direct Clue Action score, which was the best predictor of experts' grades, involved a rather elaborate system for classifying and weighting each action (5). It is very unlikely that a judge would independently formulate the details of such a system and apply it to each record he considered. Instead, the judges probably organized their task by means of successive sorting on

a few rather gross variables. The investigation described here was directed to further specification of these judging elements.[69]

When AUTOMASTS records were arranged according to the average expert grade assigned to them, two things became evident at once. First, those performances which terminated with the correct replacement were always ranked higher than the failures. Second, within the successful group of records, the shorter performances received higher grades than the longer ones.

Within the unsuccessful group of records, however, the length of a performance did not seem to be an important basis for judgment. Many reversals occurred, with a considerable number of long unsuccessful performances graded higher than shorter ones. Other indicators were tried out on the unsuccessful cases to see if satisfactory within-class ordering could be achieved. The most successful discriminator found was the Proximity score for the last quarter of each performance. This score is an average expression of how far away the technician is from the trouble.[70] Apparently, when a judge is confronted with a series of long unsuccessful records, he tends to disregard many of the details in the earlier phases and concentrates on the final portion. If

---

[69] Of course, the preferred method for investigating the question is to give the judges a "stacked deck" of specially prepared performance records in which the elements to be studied are experimentally controlled. It is anticipated that such a study will be undertaken in the future.

[70] Clue Quality or average "information value" scores for the last quarter provide about the same predictions.

a technician is proceeding well at the end of his repair attempt, then he is given a relatively high grade within the unsuccessful group. Similar "end-atmosphere" effects have been noted in other types of complex judging tasks (32).

The foregoing results can be summarized in the form of four procedural rules for classifying performances:

1. Sort the performances of a given problem into two piles--successful and unsuccessful.

2. Rank the successful performances according to the length, with the shortest ranked highest.

3. Rank the unsuccessful performances according to the Proximity Score obtained in the fourth-quarter of each performance.

4. Combine the rankings of the successful and unsuccessful performances, making certain that the highest ranked unsuccessful performance is one rank lower than the lowest ranked successful performance.

When these four rules were applied to the AUTOMASTS records, the correlation between the resulting order and the expert ordering was .95. Apparently the successive sorting procedure reproduces the expert judgments more accurately than any linear combination of scores tried to date.

Whether this particular sorting procedure would stand up in other subject samples or situations is a question that demands empirical investigation. Perhaps the main value of the present analysis is to indicate variables that deserve first consideration. Certainly the discovery and generalization of a few simple rules which predict expert ratings of quality would be of methodological and economic significance to the researcher.

Redundancy

Redundancy was defined by Weaver in his work on Information Theory in Communications (48) as that portion of a message which, if missing, would leave the message essentially complete. This concept is being applied in decision-making studies (30), and can be extended to the series of actions made by a technician in trouble shooting (18). Those actions which supply information that has been supplied already are redundant; i.e., their absence would leave the performance complete.

For present purposes eight specific classes of behavior are defined as redundant. A technician is charged with a redundancy whenever he:

A. Repeats a measurement or repeats a component replacement.

B. Takes a B+ reading at the plate and/or screen of a tube (finds the voltage normal) and then takes a B+ reading closer to the power supply.

C. Takes voltage or resistance readings at some point above ground (finds the reading is normal) and then takes the same reading nearer ground in the same circuit element.

D. Makes any reading between two points which are electrically the same and which have been measured before in the same manner (e.g., takes a voltage at two points along the receiver AVC line, then measures the voltage at a point between even though normal readings were obtained at the first two points).

E. Injects a signal between the trouble and the receiver loudspeaker, and after receiving a normal indication, injects a signal closer to the loudspeaker.

F. Injects a signal between the trouble and the radio antenna, and after receiving an abnormal indication, injects a signal closer to the antenna.

G. Takes a waveform in a stage between the trouble and the radar monitor scope and, after receiving an abnormal waveform, takes a reading closer to the monitor scope.

H. Takes a waveform between the trouble and the radar master oscillator (first stage) and, after receiving a normal waveform, takes a waveform closer to the master oscillator.

Examination of this list reveals two types of redundancy. Intrinsic redundancy, defined in A above, consists of a literal repeat. Extrinsic redundancy, defined in B through H above, involves the collection of virtually identical information by two actions which are not literal repeats of each other. The distinction between these two types of redundancy is deeper than the fact that they are based upon different sets of operational definitions. They represent qualitatively different types of mental behavior. Literal repeats can be avoided by accurate recall. Avoidance of extrinsic redundancies requires an appreciation of the inter-relationships embodied in the circuitry (plus a correct recollection of previous actions).

A single action can be redundant under two of the definitions (i.e., definition A and any one of the other definitions). For example, a man may perform an action which is classified as extrinsic on the basis of one of the definitions B through H, and then repeat that same action. The repeated act would still be classified as an extrinsic redundancy by virtue of the same definition which applied previously, but because it also fits definition A it would be an intrinsic redundancy as well. This dual classification complicates comparison between the two types of redundancy. For this report the practice was adopted of counting the

action in question as a member of both classes and adjusting the total number of actions accordingly.

The discussion of redundancy is divided into four parts. First, frequency characteristics and variability are considered. This is followed by a discussion of the most likely location of redundancies within a performance. Third, the relationships between redundancy and such factors as solution of the problems and "goodness" of performance are explored. This part appraises the validity of the assumption that a negative value should be attached to redundant behavior. The redundancy information is summarized in part four.

Frequency and Variability of Redundancies

To facilitate the treatment of redundant trouble shooting behavior, a redundancy score was developed. This score consists of the ratio of redundant actions to the total actions.[71] The use of this ratio compensates for the fact that there are greater opportunities for redundancies to occur in long performances.

Figure 22 on page 100 contains the average proportion of redundancies (average redundancy scores) occurring in the Job Sample, MASTS, and AUTOMASTS performances. About 30 per

[71] When considering the average redundancy score for a series of problems, it is not desirable to sum the scores for each performance and then divide by the number of performances as this tends to weight the shorter performances too heavily. For this reason, whenever mention is made of the average redundancy score, it will be understood that a weighted average is used.

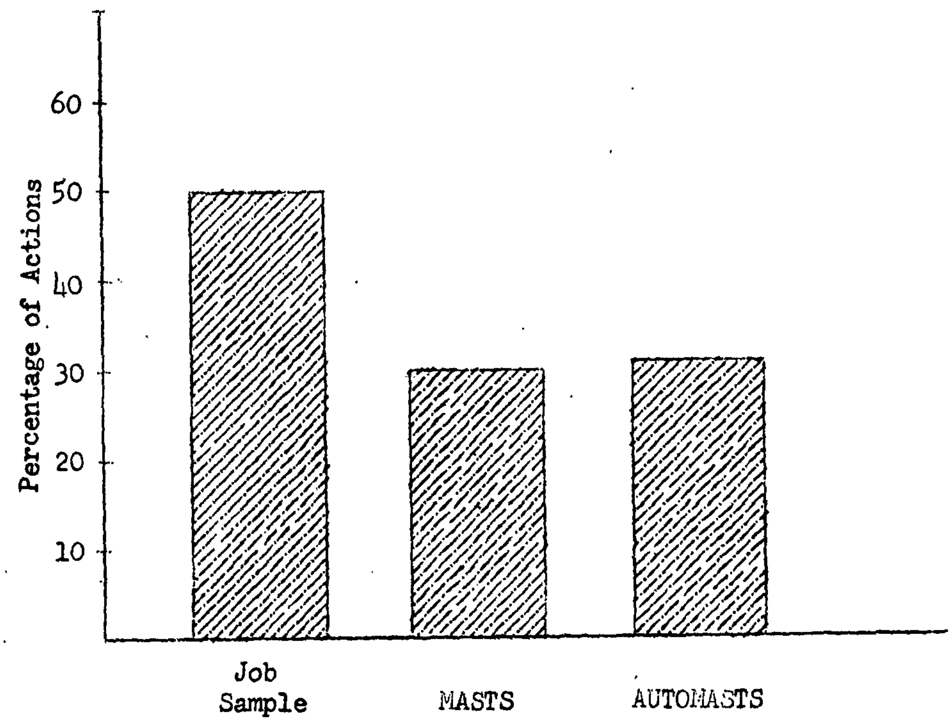


Fig. 22 Average Percentage of Redundant Actions in Job Sample, MASTS, and AUTOMASTS Performances

cent of the actions on the symbolic formats are redundant as compared with approximately 50 per cent on the Job Sample. This amounts to a substantial inter-format difference. Furthermore, it poses the question, "Why is a technician working on actual equipment

more likely to be redundant than when working on a synthetic representation of the equipment?" A consideration of the two types of redundancy suggests one reason.

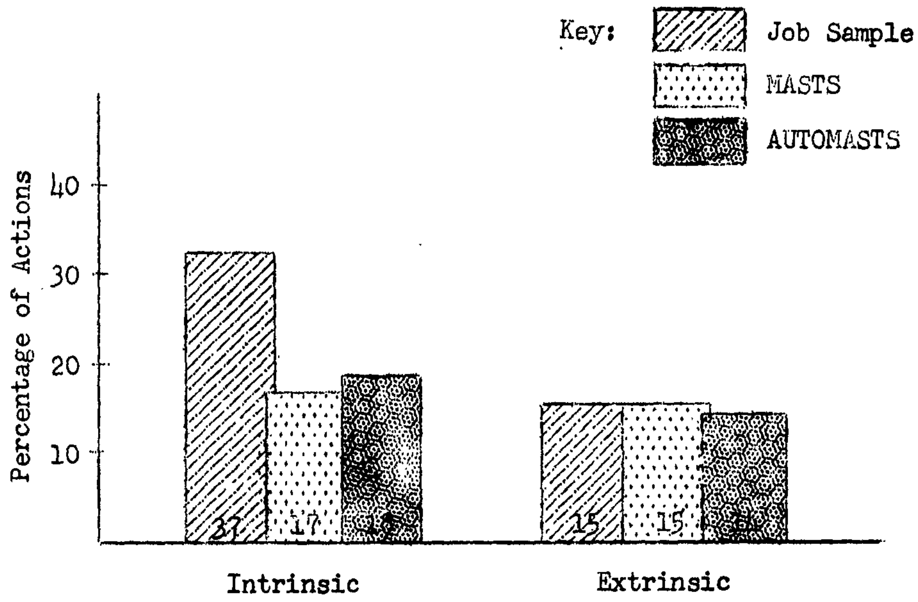


Fig. 23. Inter-format Redundancy Differences.
Heights of bars indicate the average percentage of redundant actions.

Figure 23 permits two comparisons; one among formats, and one between types of redundancy. Each percentage is weighted, as explained earlier, to compensate for repeated extrinsic redundancies. With one exception, the vertical bars are the same length (approx. 16%). A considerably larger proportion of the Job Sample performances are made up of intrinsic redundancies. This indicates that the difference (with respect to redundancies) between the Job Sample test and the symbolic tests is almost

entirely due to the relatively high incidence of repetitions on the former test.

Recognition of this fact introduces further questions. What is there about a performance on the actual equipment which leads the trouble shooter to repeat a large proportion of his activity? Why isn't this true on the symbolic formats? Several probable reasons are listed below.

> It is easier to be logical and systematic when trouble shooting a schematic diagram, than when trouble shooting a jumble of tubes, wires, and other parts. (This assumes that logical and systematic performances will contain a smaller proportion of intrinsic redundancies and that ETs fail to take full advantage of the schematics furnished them when they have some hardware to tinker with.)
>
> The test points and parts on the Job Sample equipment are not labelled, making it more difficult to remember which points have been visited. The opposite is true on the symbolic tests.
>
> Job Sample time limits are considerably longer. This allows more opportunity for redundant behavior and makes it more difficult to retain the information throughout long performances.
>
> Making a reading on the synthetic test is a more definite and simpler activity than touching a piece of wire or a contact with a test probe and then searching for a correct reading on a meter.
>
> It becomes evident in watching electronics technicians at work that they frequently do not trust their readings on test equipment. It is easy to make a poor connection with the test prod or to set the meter on the wrong scale. This mistrust causes repeated readings on the Job Sample that are not necessary on the symbolic formats where the information is certified to be correct.

There are several other characteristics of redundant behavior which should be mentioned here. Although the standard deviations for the Redundancy Scores are small (.12 - .14), the range is

very large. In both the symbolic and Job Sample formats the scores ranged from .00 to .75. This means that while some performances have no redundant actions, others are 75% redundant.

Redundant behavior is consistent. The reliability coefficient[72] across a series of problems is approximately .70 regardless of format differences or whether intrinsic and extrinsic redundancy are considered separately or together in the Redundancy Score. If an ET has a high proportion of redundancies on one problem--he'll have a high proportion on the rest.

It is of some interest to note the relative frequency with which redundancies fall under each of the operational definitions given on page 97. Figure 24 on page 104 gives a pictorial representation, from the AUTOMASTS performances, of the proportion of the redundancies located by each of the definitions. Since definition E, which only applied to the radio problems, corresponds closely to definition H (radar only), these two were combined in determining the percentages shown in the figure. The same is true of definitions F and G.

By far the most popular redundancies are intrinsic (definition A). The next most frequent (definitions E, F, G, and H) are those extrinsic redundancies which might be characterized as incorrect interpretations of the signal flow patterns. For

[72]The consistency coefficient was determined by Ebel's formula.

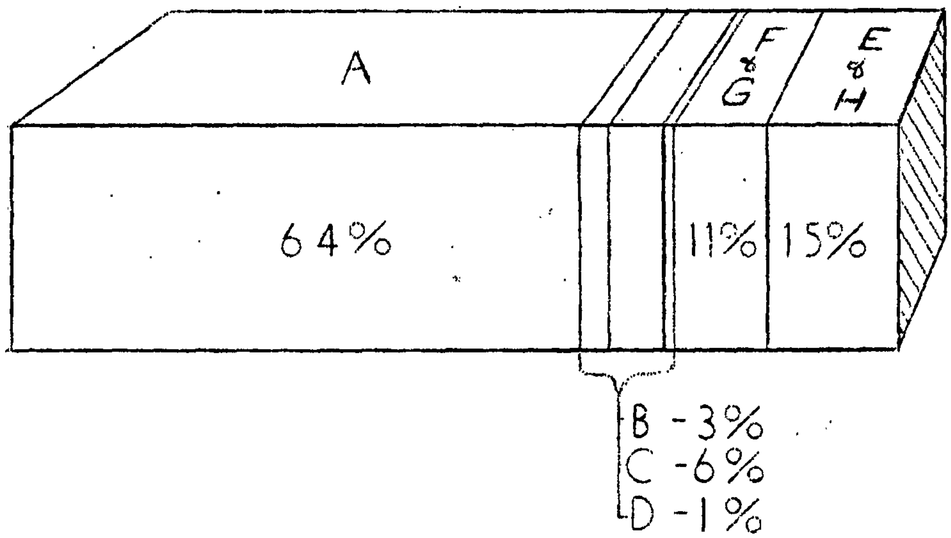


Fig. 24. Proportion of Redundancies for Each of the Redundancy Definitions.

example, definition H, the most frequent of these, involves obtaining a normal reading on the signal flow path and then, apparently misinterpreting the implications of this reading, taking further readings nearer to the source of the signal.

Location of Redundancies Within the Performance

The description of redundant behaviors should include some consideration of where they occur in the performance. The first move cannot be redundant. Other than that, there is no limiting condition imposed on the trouble shooter as to when he may make a redundant move.

Several methods were used to determine the portion of the performance where redundancies are likely to occur. The first involved splitting each performance into segments (in this case, quarters). A frequency count of the number of redundancies and number of actions in each quarter was obtained. Figure 25 shows the average redundancy score for each quarter of a combination of 422 AUTOMASTS performances.

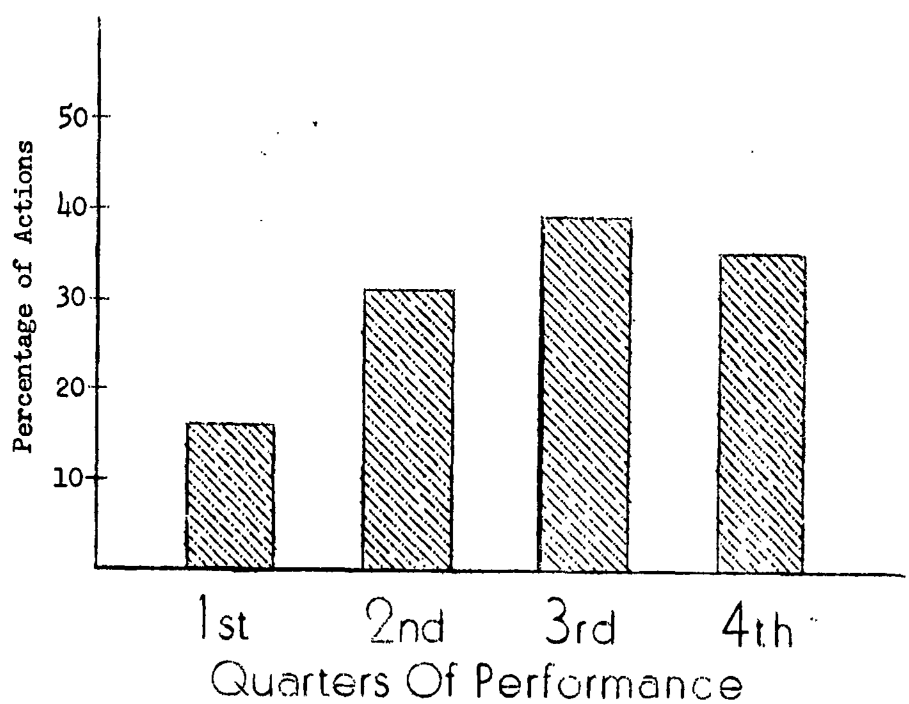


Fig. 25. Percentage of Redundant Behavior in Each Quarter of the Typical AUTOMASTS Performance.

Each of the last three quarters contains more than twice as much redundant behavior as the first quarter. Now the question arises; what about the distributions of the two forms of redundancy--intrinsic and extrinsic? Do they follow the same pattern separately as they do combined? Figure 26 shows the average proportions of moves of each type during each quarter.

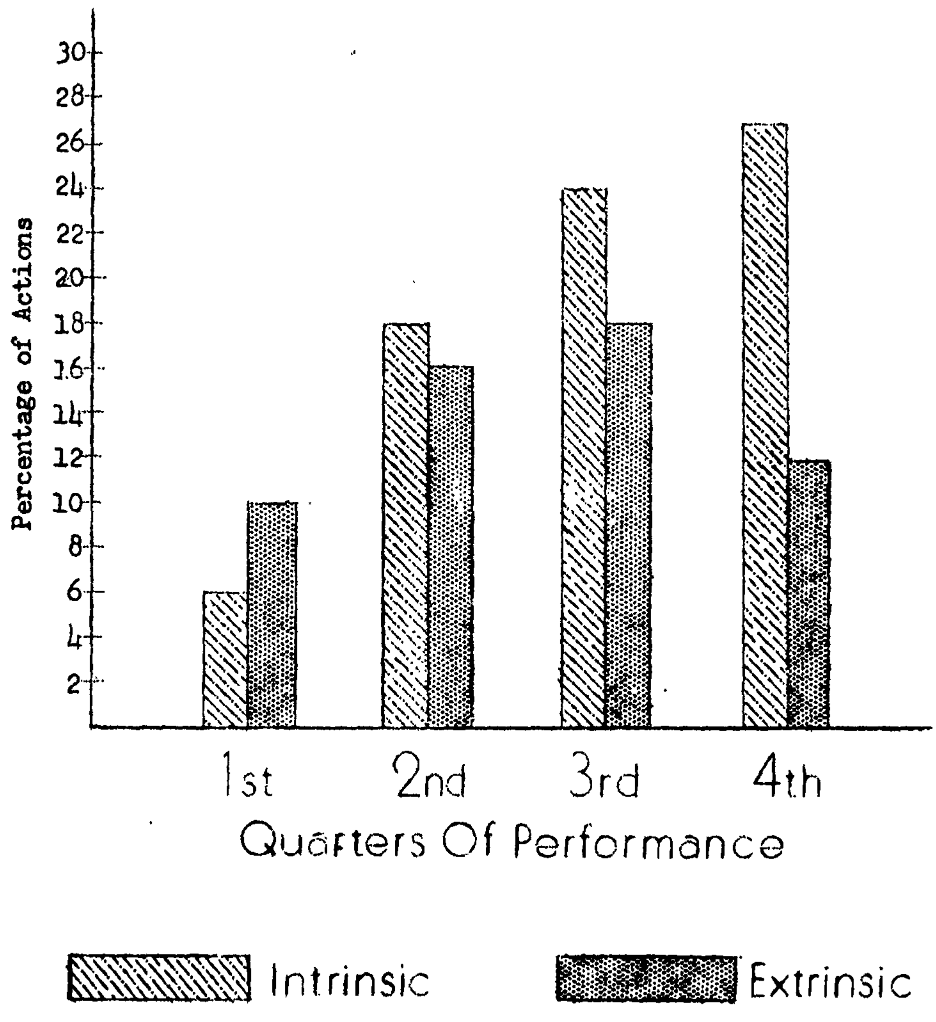


Fig. 26. A Comparison of the Types of Redundancy in Each Quarter of the Typical AUTOMASTS Performance.

It appears that extrinsic redundancies tend to hit a peak somewhat earlier in the performance than do the intrinsic redundancies. A more detailed examination of the distributions of each type of redundancy should give a better picture of the differences between them.

In order to study the separate distributions for the three types of redundancy (intrinsic, extrinsic, and dual) the trouble shooting performances were transposed to a standard 24-action base. Every redundancy was located along this base according to its relative location in the performance in which it originally occurred. The resulting distributions are presented in Figure 27 on page 108.

From an examination of the figure it is evident that the differences between the distributions are large. The distributions also differ considerably in form. The intrinsic curve and the dual curve are negatively skewed. The extrinsic distribution is symmetrical. The medians for the extrinsic, intrinsic, and dual redundancies are 10.8, 15.4, and 16.8, respectively. These findings suggest that future redundancy analyses should deal with each class of redundancy separately.

Up to now, redundant moves have been regarded as a class, separate from their identity as voltage checks, resistance measurements, etc. at particular points in the circuit. The question well might be asked: Are there any checks (and/or check points) that are repeated much more frequently than others? In examining AUTOMASTS radio performances, it is found that most

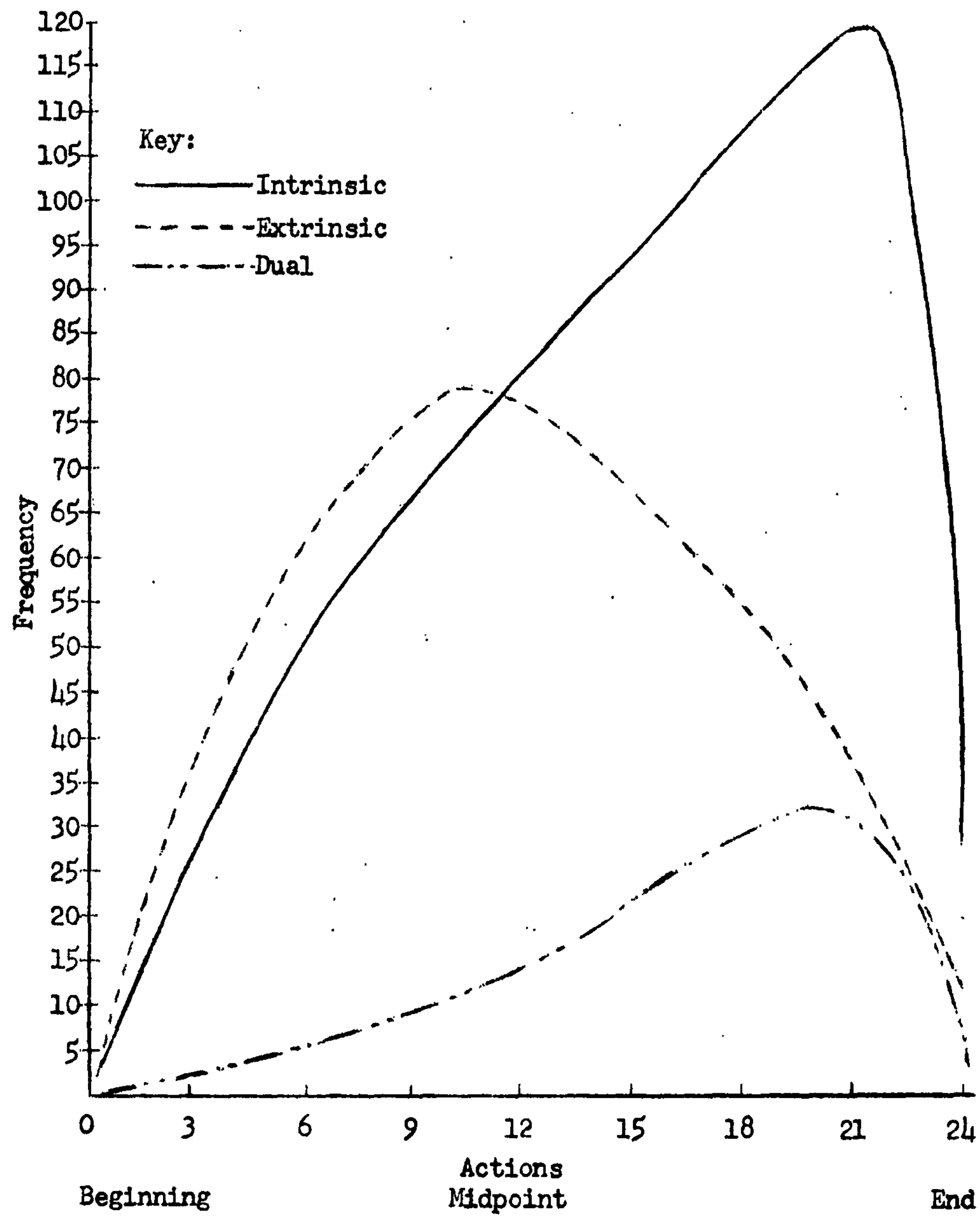


Fig. 27. Relative Frequency of Occurrence of Three Types of Redundancy as Distributed Through a Standard 24 Action Performance.

of those activities which had at least five (an arbitrary number) or more repetitions were signal generator applications, and, as would be expected, were almost always on the signal path. Almost every test point on the signal flow path of the detector stage was repeated frequently. Other than these observations, there is little to be found in a minute analysis of the actions themselves.

This completes the description of the characteristics of redundancy as far as its occurrence and general properties are concerned. The next point of interest lies in the relationship of redundancy to certain qualitative features of trouble shooting. How does this behavior relate to the ultimate success of the trouble shooter or to an independent evaluation of the quality of his performance?

Relationships Between Redundancy and Other Factors

Redundancy may be considered as either good or bad depending on the emphasis given to efficiency in evaluating trouble shooting technique. It is true that without redundant actions the performance would still be complete (by definition) and therefore more efficient. But remember, the average performance (on the AUTOMASTS) contains only 25 actions. A fourth of these are redundant, but an even greater proportion may be classified as useless actions; that is, they provide no information concerning the location of the trouble. They are probably of much less value than redundant actions and contribute much more to inefficiency. It is quite possible that those individuals who made

redundant actions, if denied the opportunity to behave in this manner, would have found it considerably more difficult to solve the problems.

In the past, investigators invariably have assumed that redundancy deserves a totally negative weight. For example, Glaser (17) includes actions of this type in a category called "inefficient checks" and weights the category negatively for scoring purposes. However, up to this time there has been little empirical evidence that all redundancy is equally bad--or for that matter, that redundancy is bad at all. It therefore is deemed appropriate to examine these trouble shooting performances to determine the effects of redundant behavior. Such an examination should indicate the justifiability of the practice of negatively weighting redundant actions in trouble shooting performance evaluations.

Two ways of testing the relationship between redundancy and success were employed. One of these is to compare the probability of success with the proportion of redundancies made on each separate problem. Table 11 on page 111 gives the correlations (point-biserial), for nine of the ten AUTOMASTS problems, between success-failure and the three types of redundancy.

Inspection of the table shows that redundancy is negatively related to success. However, the magnitude of the relationships is not great enough to warrant heavy penalties for redundancy when trouble shooting is being evaluated.

Table 11

Point-biserial Correlations of Total, Extrinsic, and Intrinsic Redundancy Scores with Success-Failure for Each of Nine AUTOMASTS Problems.

Problem	N	Type of Redundancy		
		Total	Extrinsic	Intrinsic
C102	40	-.28	-.27	-.40*
C103	41	-.53**	-.44**	-.16
C104	42	-.55**	-.54**	-.49**
C105	41	-.50**	-.57**	-.32*
C106	39	-.44**	-.59**	-.53**
C110	43	-.22	-.21	-.31*
C112	44	-.32*	-.44**	-.10
C114	45	-.41**	-.41**	-.32*
C115	45	-.54**	-.27	-.46**
Average Correlation (z-weighted)		-.42**	-.42**	-.35**

Note: Problem C111 was omitted because only one man failed this problem.

* Significant at .05 level.
** Significant at .01 level.

In addition to the problem-by-problem analysis of the relationship between redundancy and success, these two variables were compared within formats. For this comparison each man's average redundancy score was paired with the proportion of the problems that he passed on the given format. Table 12 on page 112 shows the results obtained.

Table 12

Pearson Correlations Between Weighted Redundancy Scores and Probability of Success on Job Sample, MASTS, and AUTOMASTS.

Type of Format	Type of Redundancy		
	Total	Extrinsic	Intrinsic
Job Sample (N=36)	-.40*	-.36*	-.42**
MASTS (N=36)	-.23	-.23	-.07
AUTOMASTS (N=45)	-.13	-.29	-.02

* Significant at .05 level.
** Significant at .01 level.

The Job Sample format was the only one which produced statistically significant coefficients between redundancy and success, and these are rather low. While averaging of redundancy scores across problems tends to reduce their variability (and this in itself might limit the correlations), the results seem clear. The degree to which a technician is redundant is relatively independent of the degree to which he is successful on the symbolic formats, although it is of minor importance on the Job Sample format.

A similar comparison can be made to determine the relationship between redundant behavior and experts' opinions of relative goodness of trouble shooting.[73] Table 13 on page 113 contains

[73] As explained in Report 12 of this series, each performance was rated in terms of its relative goodness and the average of three judges ratings was the Expert Judgment score for that performance.

the correlations between the expert judgments and the three types of redundancy.

Table 13

Rank-order Correlations of Total, Extrinsic, and Intrinsic Redundancy Scores with Expert Judgments of Performance Quality for Each of Ten AUTOMASTS Problems.

Problem	N	Type of Redundancy		
		Total	Extrinsic	Intrinsic
C102	40	-.48**	-.38*	-.42**
C103	41	-.51**	-.72**	-.22
C104	42	-.58**	-.58**	-.45**
C105	41	-.67**	-.74**	-.51**
C106	39	-.74**	-.77**	-.51**
C110	43	-.56**	-.60**	-.55**
C111	44	-.76**	-.72**	-.68**
C112	44	-.33*	-.44**	-.19
C114	45	-.50**	-.57**	-.36*
C115	45	-.32*	-.48**	-.18
Average Correlation (z-weighted)		-.56**	-.62**	-.42**

* Significant at .05 level.
** Significant at .01 level.

There appears to be a substantial negative relationship between total redundancy and the quality of performance assigned by experts. The relationship is even higher between extrinsic

redundancy and performance quality. On this basis it seems probable that the experts gave some weight (negative) to extrinsic redundancy.

As in the case of the comparison with success, it is possible to discuss the relationship between redundancy and expert judgments of the combined ten performances for each man. When the average redundancy score (of each type) is correlated with the average expert judgment (by man) the resulting coefficients (-.45, -.63, and -.29, for the total, extrinsic, and intrinsic, respectively) indicate a clear-cut difference between extrinsic and intrinsic redundancy.[74]

It seemed likely that the variability among coefficients of this type was related in some degree to the difficulty levels of the problems. This was investigated by ranking the problems in terms of their difficulty levels and comparing this with the rank order of the correlations shown in Tables 11 and 13. The coefficients between problem difficulty and the correlation of redundancy and problem success were not statistically significant. In other words, the relationship between redundancy and success is independent of problem difficulty. However, the experts' judgments do seem to be influenced by problem difficulty. Significant relationships (beyond the .05 level) were found between the

[74]While the extrinsic coefficient is statistically significant beyond the .01 level, the intrinsic coefficient is not significant.

difficulty levels of the problems and expert-redundancy correlations. This probably indicates that as the difficulty level increases, the experts are less able to take into consideration such factors as redundancy in making their judgments.

Summary of Redundancy Analyses

The concept of redundancy was related to trouble shooting data by eight operational definitions, each representing an activity which would yield information already obtained in previous action sequences. Redundant behavior was also divided into two overlapping categories--Extrinsic Redundancy (non-repetitive redundancy) and Intrinsic Redundancy (repetitive redundancy). The overlap was taken into account by combining the overlapping cases with both categories when comparisons between the two were made.

Considering redundancy as a whole, about 30 per cent of the average performance is redundant on the symbolic formats. This increases to about 50 per cent on the Job Sample performances. The increase is attributable to the intrinsic type of redundancy and may be due to the technician's lack of confidence in using test equipment and the difficulties associated with working in the maze of components, wires, etc. of the actual equipment.

Redundant behavior is widely variable, ranging from 0 to 75 per cent of the performances. However, technicians are found to be consistent from problem to problem, some being consistently more redundant than others.

About four times as many repetitive (intrinsic) redundancies occur as do those of any of the other seven definitions. Of the extrinsic redundancies, the most frequently occurring are those which might be categorized as misinterpretations of the implications of previous actions.

A disproportionate number of the redundant actions occur in the last three-fourths of the performance. Extrinsic redundancies are most likely to appear in the middle of the performance, whereas intrinsic redundancies are most likely to happen near the end.

To aid in the evaluation of redundant behavior as it relates to good trouble shooting technique, comparisons were made between redundancy and success, and between redundancy and judged "goodness" of the performances. The results indicate a small, significant negative relationship between redundancy and success. It appears to be too small to justify the large negative weight given to this behavior by most investigators.

On the other hand, the judgments of the experts on performance quality indicate a substantial relationship between "good" performances and a minimum extrinsic redundancy score.

Errors

With the freedom of action given a trouble shooter, it is inevitable that some of his checks will involve improper application of test equipment or generally unsatisfactory work practices. Such activities are called errors. This portion of the

report is devoted to a brief treatment of the frequency and consequences of different classes of errors.

The Job Sample situation provides the best opportunity for really serious errors to occur, for here the technician can actually endanger himself and can thoroughly foul up the equipment. A search of Job Sample performance records shows that a serious error occurred, on the average, about once every four performances. These <u>major errors</u> can be grouped into three classes. The first class involves the accidental receipt of an electric shock by the technician as he worked. About 15 per cent of the major errors are of this type. From the records, it is not always possible to discern the exact conditions under which these shocks occurred. There are, however, several instances where the technician was holding a test probe at the time of shock and accidentally contacted the uninsulated part of a test prod, as well as other instances when he was just "fooling around" in the gear with his hands. As far as can be determined, all shocks were the result of personal carelessness. In at least one case, the burn from a shock was severe enough to occasion first aid treatment.

The second class of major errors, consisting of about 60 per cent of the total, involves misuse of the multimeter. It includes setting the meter on the wrong numerical scale (e.g., 10X instead of 100X), setting the meter to the wrong type of measurement (e.g., ACV instead of DCV), and purposely ohming check points in a live set. Such activities are liable to wreck a

multimeter (--and sometimes did). Over half the subjects made errors of this type at one time or another during their Job Sample problem series.

The third class includes putting a removed part back in the wrong place, reversing leads, or hooking up the signal generator or oscilloscope in an unsatisfactory manner. Occasionally, these errors led to radical changes in the operating characteristics of the set, with subsequent sparking, short-circuiting, and blown fuses.

The consequences of some major errors are widespread. For example, if a technician puts a part back in a place where it doesn't belong, he may not discover his mistake and the rest of his work on that problem will be ineffective. Or he may observe his faulty replacement and correct it without serious effects on his later activity. But in the statistical sense, major errors clearly have a negative influence on likelihood of success. Of those performances which contained major errors, only 24 per cent ended in success. This figure is in sharp contrast with an overall solution likelihood of 60 per cent for the problems examined. For 36 Job Sample subjects, the number of major errors made by the subjects is significantly correlated ($r = -.55$) with the total number of Job Sample problems solved and also with the total number of Job Sample and MASTS problems solved ($r = -.55$). On the other hand, correlations with ratings of shipboard trouble shooting ability were insignificant ($r = -.25$). When one considers these facts, and the personal and practical inconvenience

caused by major errors, it is clear that errors of this type are a deterrent to effective trouble shooting. Definite steps should be taken to avoid them.

Apart from major errors, there are check behaviors which represent improper use of test equipment, in the sense that the test instrument is applied to inappropriate points. These behaviors are designated as minor errors. Though no danger or damage is involved, such checks cannot possibly yield any useful information concerning the operating conditions of the gear being checked. On the radio, they consist of voltage and resistance readings at grounded points or at places such as the antenna where voltages are too small to be detectable, and signal injections into the AVC line, cathodes, screen grids, and power supply. Similar voltage and resistance minor errors are found on the radar, as well as inappropriate waveform checks at grounded and power supply points.

Minor errors are quite frequent in Job Sample trouble shooting; all subjects made some. The average performance has two or three, and some performances have over twenty.[75] Unlike major errors, however, minor errors seem to bear little relationship to likelihood of success; correlations with number of Job Sample problems solved, number of Job Sample and MASTS problems solved, and

[75]The correlation between major and minor errors for 36 Job Sample subjects was .43, significant at the .05 level.

shipboard proficiency ratings were -.13, -.10, and .12, respectively. All these coefficients are statistically negligible for N = 36. Apparently a technician can make minor errors and get away with it.

In the symbolic formats also, technicians could make minor errors.[76] The frequency was slightly less than in the Job Sample, averaging between one and two for a performance. The same kinds of improper signal, waveform, voltage, and resistance readings were observed as in the Job Sample minor errors. Correlations between minor errors and other indices of trouble shooting effectiveness are low; on AUTOMASTS data, correlations between the error score and number solved, average time spent on a problem, and expert judgment score were -.25, +.27, and -.44, respectively. Only the last of these coefficients is significant. These results, of course, are in general accordance with minor error analysis on Job Sample data.

In summary, major errors, which may endanger the man and the equipment, are due mostly to carelessness of the technician in his work practices. They are relatively infrequent, but when they do occur they reduce the technician's chances of success. It also appears that the better technicians make fewer major errors. Minor errors, on the other hand, are rather frequent, but they seem to make little difference in the outcome of a

---

[76] Out of a total of 600 actions, each AUTOMAST problem disc has about 100 different actions which are classed as minor errors.

particular problem or in other indices of trouble shooting proficiency.

Trouble Shooting and Problem Solving

The study of solution processes is an active discipline within experimental psychology. At least some trouble shooting behavior appears to be of the problem solving type. Yet our analyses to this point have not seemed to rely very heavily on general problem solving models or findings. In this connection, therefore, it appears appropriate to clarify some of the relations between trouble shooting and problem solving and to consider, briefly, a few parallels between these two spheres of investigation. The present treatment is addressed primarily to those who are experienced in electronics maintenance matters and who are not especially familiar with the methods and results of problem solving research.

Though several pages could be spent on definitions, we shall be concise. Trouble shooting includes any activity which is directed expressly to the correction of certain classes of malfunctions. A problem exists when a goal is recognized but the meaning or route to the goal is not immediately clear. On the basis of these definitions, all trouble shooting is not problem solving, since for many trouble shooters the path to the goal is quite clear-cut and routinized. Nevertheless a considerable overlap exists. We can say that insofar as a trouble shooting attempt

contains exploration of the situation to determine the crucial elements, alternative approaches, and the relationships between them, then it involves the type of behavior usually called problem solving.[77] Probably most of the situations analyzed in this report qualify as being genuine problems. Granting this, what can the general psychology of intellective processes contribute to an understanding of trouble shooting?

Many students of problem solving have attempted to formulate a standard outline of the solution processes (11, 31, 47). In the usual schema of this sort, several phases are postulated. First there is "recognition of the problem," or awareness that a perplexed situation exists. Next, some "orienting observations" are made; they contribute to further definition of the problem and suggest possible approaches or manipulations. These are followed by a "search phase" in which critical data can be obtained. More or less systematic search leads eventually to one or a series of "hypotheses." The presumed consequences of these

---

[77] There are many procedural differences between a trouble shooting task and the conventional problems used in the laboratory. Usually the laboratory or "academic" study of thinking is designed around a specialized type of problem which can be separated readily into a few controllable aspects. The experimenter then can manipulate variables rather precisely. He also can arrange the situation so that knowledge and experience variables can be ruled out, and so that the quality of performance or the presence of certain behavioral consequences may be evaluated simply. These conditions are seldom imposed on trouble shooting data of the type considered in this report; the trouble shooting problems are complex and often lengthy, the range of analysis is more broadly conceived and therefore less clear-cut on specific issues, and experience differences undoubtedly exert a tremendous influence.

hypotheses, if in fact they do indicate the sources of difficulty, are then (ideationally) "elaborated" in relation to the other problem data. Finally, "verification" or validation of the hypothesis takes place.

Sometimes an experimental situation can be arranged so that these phases are separately discernible. But how useful are such general schemas when applied to our data? In the present study, "recognition of the problem" and identification of gross trouble symptoms were excluded as important variables by providing each technician with explicit symptoms.[78] As to "orienting observations", it is questionable whether this kind of behavior is productive, if one takes it to include manipulating the front panel controls and switches. MASTS and AUTOMASTS subjects do not find such activities necessary, though they have ample opportunity to make orienting checks. The search phase in a trouble locating task is not as sharply demarcated as the logically derived schemas imply it is; the technician displays some kind of search at nearly every point in a performance. Perhaps search is too inclusive a term, and has to be broken down further.[79] Instead of a clear-cut emergence of a series of definite hypotheses, each of which

---

[78] Saupe (41) found that ability to restate symptoms was not a discriminator between good and poor radar mechnics.

[79] The schema developed in Section III represents one attempt in this direction. It was useful to separate search behavior at the beginning of a performance from search behavior that occurred inside a stage or following a replacement.

is individually "elaborated", the trouble shooting protocols indicate that something like a "hypothesis zone" is more typical. This "zone" becomes progressively differentiated as the technician proceeds from localizing to isolating behavior. "Elaboration" is often fragmentary or superficial. Apparently the technicians believe the procedures they follow preclude the necessity of a comprehensive reconsideration of the whole problem before every component replacement. These observations suggest that although general phase models of problem solving can be imposed on trouble shooting data, such models are not apt to be enlightening until the phases, and the relationships among them, can be specified under various conditions. Models specifically aimed at the trouble shooting situation are more likely to provide an improved framework.

The concept of "set" is important in describing many kinds of behavior, and it is often referred to as a process variable in problem solving. The presence of a set implies that the subject behaves in a prepared, organized way--that he attends to certain aspects of the situation and ignores others and that he has made some of his procedural decisions in advance. Many experiments show that the subject who is properly set behaves more predictably, smoothly, and effectively; preparation may be half the battle (32, 46). Many other studies show the negative influences of fixed assumptions about a problem (34, 38, 44). Whether facilitating or inhibitory, the dynamics of set are

likely to be important when a human subject is faced with a complex situation. Almost any behavior involves some kind of selectivity, so that it is often difficult to decide whether a set is operating or not. The clearest demonstrations of sets are obtained experimentally, for here the set can be attributed to instructions or other special arrangements on the input side, and its effects and duration traced to response characteristics on the output side (32).

Sets can also be inferred if the problem situation is suitable for certain types of behavior patterns to be discerned. In the present data, for example, all the technicians had some sort of a "narrowing" or localizing concept in terms of equipment units larger than a single component; they preferred to work along the main signal or pulse paths; they usually ignored AC voltages. At a more specific level, some of the behavior patterns can be said to exemplify rather detailed and highly organized sets. The extended treatment of IAS methods, treated in Section III, provides ample documentation of this point.

If such behavior patterns are viewed as manifestations of sets, what is gained thereby? From the research point of view there are several implications. There is considerable laboratory information on the way that subjects learn and utilize sets. For example, it is desirable for a subject to have active participation in applying the set at the time it is to be learned (15). Also available from the laboratory are indications that disadvantageous sets can be "broken" (32, 46), that <u>susceptibility</u> to set is

distinct from ability to overcome set (24), that rigid adherence to a set is accentuated by highly speeded working conditions (34), and that the most flexible problem solver is the one who has a large stock of applicable sets and special training in shifting from one to the other (32, 40). All these facts are subject to verification and extension with trouble shooting materials. From the technician's standpoint, it may be desirable for him to realize some of the effects and limitations of sets, the dangers of adhering to a pet approach, and so on. Perhaps a good practical rule would be for the technician to verify his localization by more than one approach, if possible. This would at least provide him with more than one "slant" on the problem data and might reduce premature localizations and replacements. The within-stage search patterns, in particular, seem to be rather poorly structured, and a definite statement of the kinds of sets suitable for different stages might be very useful.

The trouble shooter's behavior is always subject to the effects of transfer. Such factors as his general electronics knowledge, and his facility with test equipment are all developed on the basis of previous experiences. Obviously the transfer effects, just as set effects, may be either positive or negative. Based on laboratory evidence, positive transfer is more likely when output symptoms and causes are rather similar (32), when the performance in one situation is based on an integrated understanding of principles rather than a recall of

isolated facts (46, 49), and when the earlier task has been critically reviewed by the solver himself and by others (32). General attitudes of deliberation, and caution in making inferences also would be expected to provide improvement and extend across situations. Such attitudes themselves are trainable (32). Negative transfer effects are most likely when two troubles have the same output symptoms and involve the same types of circuits but require different kinds of detailed tracing, or when the underlying principles governing the two troubles are "reversed" or opposite. As to general transfer to all types of problems, the experimental evidence suggests that it may be quite small.

The extent to which problem solving is a "trial and error" or an "insightful" process has been argued for a long time. The term trial-and-error was often utilized in animal psychology to designate sequences of blind, thrashing movements. There is no need to attach such a negative evaluation to trial-and-error in human problem solving. Some kind of trial-and-error is almost inevitable in a real problem; the subject may have to make tentative efforts to gain information and structure the task. It often appears that he is acting without any clear notion of exactly what the crucial data are. But it must be emphasized again that such trial-and-error behavior in the trouble shooting situation is not random. Even when a technician jumped around from stage to stage, he was likely to choose the more informative check points, not just any landing place. And he

was not apt to inject signals into the on-off switch or pull any part that happened to be handy. As has already been noted, unsystematic behavior occurred rather frequently within the confines of a stage. In many of these instances, the technician's view appears to be that, once he has localized, he can "afford" to be unsystematic.

An insightful solution involves the understanding of the functional relations that exist (49). Once this understanding is penetrating and comprehensive enough, certain "forces" or "tensions" will operate to span the gap between the "givens" of a problem and the desired solution. Generally there will be some "restructuring" or "recentering" of the main elements of the problem. In one sense, every time the technician recognizes the implications of a check reading he achieves some reorientation to the problem, some new awareness. The classical type of comprehensive "recentering", however, was observed rather infrequently in our protocols. There were occasional examples where a technician was "plodding" along and suddenly seemed to realize that a particular sequence of moves could provide him with significant information. Rapid accomplishment of the sequence then led to a considerable saving of time.[80] But in

[80]Outstanding instances of restructuring were observed in a problem which had low plate voltages throughout the set. Each stage was connected to a common B+ voltage bus bar. Some subjects saw that, by hooking the voltmeter to the plate of a tube and then removing one stage at a time, the offending stage would be identified when a removal caused plate voltage to return to normal.

the main, a distinctive characteristic of trouble shooting behavior is the partial degree to which such formulations are made and the tentative nature with which they are followed up. It appears that the technicians are not very confident. On our most difficult and challenging problems, it was indeed necessary to achieve some reorganization of the information complex. But the recentering could occur only after a considerable amount of regular check sequences had been made, and it was seldom completely novel, sudden, or comprehensive.

From all this it appears that trial-and-error and insight are complementary to each other. It may be, as Sargent's experiments suggest (39), that the difficulty level of a problem is related to the amount of restructuring required, and that restructuring is likely to be a highly unstable phenomenon if it is attempted too early. The technician should recognize that his insights are more apt to be useful when they are based on a solid base of problem data.

To what extent does good trouble shooting depend on the technician's basic abilities such as reasoning? A few factors such as verbal comprehension, general reasoning, and numerical facility have small positive correlations with trouble shooting performance (9, 21).[81] Spatial visualization and symbolic

[81] These correlations undoubtedly are reduced somewhat because the technicians have already been selected for training on the basis of verbal and numerical ability.

reasoning also appear to be promising variables for some kinds of performance prediction. But much higher predictions are yielded by job knowledge tests which assess strictly technical information such as theory and knowledge of operating characteristics. This result, which holds true for all kinds of corrective maintenance jobs (29, 43), again emphasizes that for the usual technician population, technical knowledge may far outweigh basic ability factors.

What tips or hints can the laboratory work on problem solving contribute to the trouble shooter? Several rules can be stated that have been found effective in one experiment or another. When paraphrased in trouble shooter's terms, they might appear as follows:

1. Be sure that schematics, manuals, tools and test equipment are close at hand.

2. If more than one person is working on the trouble, designate one person as the task director.

3. Specify, as completely as possible, the gross symptoms; try to utilize front panel indications as clues before going into the equipment.

4. Break up the task into definite units.

5. Formulate an instrumental plan that permits rapid elimination of certain factors in the situation. This may involve only a few key checks. Once elimination has been achieved, concentrate on the remaining materials.

6. Insofar as possible, decide in advance how far a given line of attack is to be pursued.

7. Adopt a suitable working pace, but do not adhere to it slavishly.

8. Record, by means of writing or a check list, the relevant data on the trouble as the search proceeds.

9. Label or otherwise arrange all removed parts so that they can be identified positively.

10. If verification or confirmation of data must be accomplished, do it immediately so that continual and wasteful back-tracking will be minimized.

11. If the available evidence points to one source of difficulty, decide what other evidence would validate the interpretation, and try to obtain it.

12. Always frame the problem data in terms of functional relationships in the equipment; be hesitant of relying on memory of similar difficulties in the past.

13. Do not allow the mere accessibility or convenience of parts of the equipment to determine the search plan.

14. Try to isolate large blocks or units for checks of gross functioning.

15. If possible, try to check a part in the equipment before removing it.

16. Be wary of making radical alignment or other changes in the equipment that would make it difficult to restore the equipment to its original condition.

17. If a chain of data-gathering and reasoning is unsuccessful, formulate the assumptions on which the inference is based, and see if alternative assumptions are tenable.

18. When a "dead zone" or unproductive phase is reached, stop working and adopt a critical "spectator" attitude before going into the task again.

19. Be prepared to shift the approach to the problem and come at it from a different angle.

20. Resist the temptation to remove components at random within a section of the equipment, unless the components are very few in number and no cross checks are possible without removal.

Most of the above principles are obvious enough, and every technician follows them to some extent. It undoubtedly would

be of value, however, for such principles to be recognized explicitly and organized in technical training. The great need is to introduce a more systematic approach throughout; the observer cannot miss the overall impression of vagueness and uncertainty with which the average ET formulates his trouble locating task and proceeds with a search program. There are already indications that specially designed guide materials can greatly aid in systematizing the work. Even simple trouble locator sheets may be of substantial benefit (1).

In summary, our brief survey has shown that some methods and results of laboratory problem solving research can be identified in the corrective maintenance situation. The same general models apply, and concepts such as set, flexibility, transfer from one problem to another, and task restructuring are common and useful to both fields. In its present state, the main positive contribution of laboratory studies will be a stock of suggestions which can be tried out in the trouble shooting situation. Some of these possibilities have been mentioned above and there are many others. For example, it might be profitable to investigate the validity of the "hints to the problem solver" listed above, and the extent to which they can be implemented practically. The laboratory results help to point out features of problem solving on which the training and supervisory official, and the technician himself, should focus their attention.

SECTION V. TEMPORAL CHARACTERISTICS OF TROUBLE SHOOTING

Up to this point little has been said in this report regarding the time aspects of the performance records. It will be recalled, however, that in the Job Sample and MASTS tests minute-by-minute time entries were made, and in the AUTOMASTS test the duration of each performance was noted.

These timing procedures were carried out in order to provide data from which, it was hoped, important inferences could be drawn about the trouble shooting process. It was felt that the manner in which a technician distributes his actions along a time base might provide important clues as to his general structuring of the problem. In addition, the time factor has scalar properties which make it useful for analytic purposes. Since time is rather easily measured and interpreted, it can provide a practical basis for organizing and inter relating diverse variables.

Timed versions of these tests provide suitable data for analysis since each problem is a standardized task to be completed as quickly as possible. Therefore, the subject's speed is determined chiefly by his interactions with the intricacies of the electronic circuitry; it is not influenced by distractions, interruptions, availability of spare parts, or other irrelevant factors.

The general objective of this analysis is to relate temporal characteristics to other performance variables, in a way that will be helpful in integrating some of the analyses in previous sections. While interpreting these trouble shooting variables within a time

framework we shall be particularly interested in evaluating the extent to which time measures are useful for predicting individual differences.

Basic Treatments of Time Data

The analysis which will be presented is derived completely from the time entries made by the observers as they recorded the actions of the technicians in the trouble shooting tests. The observers were instructed to write down the time on the form provided at least every minute in the Job Sample and MASTS tests and at the end of the performance in the AUTOMASTS test.

With these performance records it becomes possible to consider time as a variable in four general ways. First, we can deal with the time to a particular point in the record, including of course, time to the end of the record.

Secondly, we can center our attention on various rate phenomena, expressed either as actions per unit of time or as time per action. These rates can be presented for the total performance on a problem, or for specified segments of the performance. In this report all such rates are expressed in terms of actions per minute.

Thirdly, we can attempt to point out significant changes in tempo within a performance and to relate these to the technician's behavior.

Fourthly, we can abstract out particular behaviors, wherever

they may occur in a performance record, and attempt to relate these to time factors.

Time Analysis of Entire Performances

Since time for solution and overall speed of working have been frequently mentioned as criteria of goodness in trouble shooting[82], these measures will be dealt with first, as if no "internal" timing of the records were available.

Duration of Entire Performances

In Table 14 on page 136 are shown the means and standard deviations for performance times of various groups, along with the number of performances in each group. In connection with this table it should be recalled that the time limit for each problem on the Job Sample test was 35 minutes; for the MASTS test, 20 minutes; and for the AUTOMASTS test, 10 minutes for the radio problems, and 15 minutes for the radar problems. Thus, it can be seen that the figures of most interest are the times for successful performances, since the unsuccessful performance times (which closely approximate the time limits) are incorporated in the total group times.

For the Job Sample test the mean solution time was 12.06 minutes; for MASTS, 7.44 minutes; and for AUTOMASTS, 6.47 minutes.

[82]Several investigations (7, 14, 16, 25) have used some such time factor in evaluating trouble shooting.

Table 14

Total Performance Time Summary Data
(All times are expressed in minutes)

Data Group	Number of Subjects	Number of Problems Used	Successful Performances			All Performances		
			Number	Mean Time	S.D. of Time	Number	* Mean Time	S.D. of Time
Job Sample Radio	36	12	163	11.14	8.24	216	16.80	12.28
Job Sample Radar	36	12	121	13.29	9.54	216	22.51	12.73
Job Sample Total	36	24	284	12.06	8.88	432	19.65	12.83
MASTS Radio	36	12	126	7.29	5.36	216	12.54	7.45
MASTS Radar	36	12	116	7.61	5.62	216	13.34	7.42
MASTS Total	36	24	242	7.44	5.49	432	12.94	7.45
AUTOMASTS Radio	45	5	128	5.77	2.38	225	7.59	2.76
AUTOMASTS Radar	45	5	113	7.27	3.81	225	11.12	4.72
AUTOMASTS Total	45	10	241	6.47	3.22	450	9.35	4.25
All Radio Problems	81	13	417	8.33	6.51	657	12.24	9.20
All Radar Problems	81	12	350	9.46	7.37	657	15.59	10.16
All Problems	81	25	767	8.84	6.94	1314	13.92	9.84

*Cut-off times were employed on all formats. In computing average total times the exact times for unsuccessful performances were used, which allows for the fact that a few unsuccessful performers gave up before the time limit was called.

It can be seen that the magnitudes of these mean solution times follow the same order as the time limits, and that in each format the mean solution time for the radio problems is less than for the radar problems. The standard deviations of the solution times are relatively large and show perfect positive correlation with their means. The mean solution times generally fall within the middle third of the time allowances, indicating that the selected time limits are probably not too severe.[83]

Action Rates for Entire Performances

Referring to Table 15 on page 138 we see a summary of the action rate (actions per minute) data for all major subgroups. In this table the means and standard deviations of action rates for successful and unsuccessful performances, separately and combined, are shown. These means are computed by dividing the sum of rates by the number of performances, thereby weighting long and short performances equally. Although the differences between formats due to time limit differences may be ignored for purposes of rate comparisons, one other distinction between formats should be considered. For the three formats there are important differences in the definitions of an action.

Whereas in the AUTOMASTS test six different types of actions were recorded, in the MASTS test eight different types were used, and in the Job Sample test, 22 different types. These differences

[83] Further data bearing on the suitability of the time limits, based on nine selected MASTS problems are shown on page 157.

Table 15

Action Rate Summary Data
(All rates are expressed as actions per minute)

Data Group	No. of Subjects	No. of Problems Used	Successful Performances			Unsuccessful Performances			All Performances		
			Number	Mean Rate	S.D. of Rate	Number	Mean Rate	S.D. of Rate	Number	Mean Rate	S.D. of Rate
Job Sample Radio	36	12	163	2.97	1.66	53	2.08	.70	216	2.75	1.53
Job Sample Radar	36	12	121	2.50	.97	95	2.07	.66	216	2.31	.87
Job Sample Total	36	24	284	2.77	1.43	148	2.07	.67	432	2.53	1.27
MASTS Radio	36	12	126	3.03	1.14	90	2.05	.74	216	2.62	1.10
MASTS Radar	36	12	116	2.89	1.04	100	2.22	.71	216	2.58	.96
MASTS Total	36	24	242	2.96	1.09	190	2.14	.73	432	2.60	1.03
AUTOMASTS Radio	45	5	128	2.97	1.07	97	2.74	1.08	225	2.87	1.08
AUTOMASTS Radar	45	5	113	2.61	1.09	112	2.26	.93	225	2.43	1.03
AUTOMASTS Total	45	10	241	2.80	1.10	209	2.48	1.03	450	2.65	1.08
All Radio Problems	81	13	417	2.99	1.35	240	2.33	.94	657	2.75	1.26
All Radar Problems	81	12	350	2.66	1.05	307	2.19	.79	657	2.44	.96
All Problems	81	25	767	2.84	1.23	547	2.25	.86	1314	2.60	1.13

in definitions arose in the development of the three formats to satisfy particular needs which were not necessarily related to the present time analysis. This circumstance makes the cross-comparison of formats more difficult wherever the number of actions is concerned, but with the introduction of certain qualifications most of these comparisons are still meaningful.

The Job Sample coding of actions gives the most complete account of trouble shooting activities. The observers recorded nearly every activity they could see, including listening, sniffing, looking at schematic diagram or parts booklets, adjusting meters, turning switch on or off, etc. Obviously not all of these behaviors should be considered actions in a scoring scheme, although these notations along with the marginal comments which were recorded on the data form during the test were helpful in interpreting the test protocols.

For the purposes of calculating rates it was decided to exclude from the action count those activities which seemed useful merely for facilitating a check or which were rather personalized behavior. Consequently the definition of an action was taken as:

> A specific test (a) to obtain instrument indications, (b) to find out what observable changes in the operation of the gear occur in responses to performing some operation upon it or from replacing a component, or (c) to determine the effects of injecting standard signals.

Using this definition, basically the same action categories were used in the Job Sample test as were used in the MASTS test. Nonetheless, there were still some differences in the activities which were counted as actions in the various formats. For example,

in the Job Sample test a component replacement was coded twice--once when first removed, and once when replaced with a new (or the same) component. Also, test instrument readings had to be made with one probe at ground in the MASTS and AUTOMASTS formats, while Job Sample readings could be taken between any two check points. Differences such as these are generally inconsequential, but where the definitional differences seem important the formats are treated separately.[84]

The overall action rate for Job Sample problems is shown in Table 15 as 2.53 actions per minute, for all MASTS problems as 2.60, and for AUTOMASTS as 2.65. From this table it also appears that, based on all performances which fit each definition, rates are higher on radio than on radar problems, on successful than on unsuccessful performances, and on successively more "symbolic" formats. These relationships and other specific hypotheses about overall action rate will be examined in the paragraphs which follow.

Consistency of Each Subject's Action Rates

A number of questions immediately arise regarding the use of action rates. The first one that will be considered concerns the consistency of the individual trouble shooter's rates. Are there stable differences among subject's rates from problem to problem? To what extent are we justified in talking about "fast" or "slow" trouble shooters? Stated in another way, is it true that the

[84] A correlation of .96 was found between total number of actions on Job Sample problems (4 radio and 5 radar) coded by Job Sample definitions versus AUTOMASTS definitions for 144 performances.

subjects tend to maintain the same relative rankings as to speed over a series of problems, when compared to others performing the same tasks? In testing this consistency it was assumed that format or equipment differences might affect individual action rates differentially or otherwise obscure the results. For this reason separate analyses were made by format and by equipment, using an analysis of variance approach.[85]

Table 16 on page 142 presents data on the question of inter-problem consistency of action rates for individuals in six homogeneous groups. For each of these groups coefficients of consistency were computed by two different analysis of variance formulas. The formulas used in computing these coefficients and testing their significance, along with complete analysis of variance tables, are shown in Table I of Appendix C.

The first column of coefficients shown in Table 16 represents the average inter-problem correlation for all possible pairs of

---

[85]Before discussing these results, it should be pointed out that there is some difficulty in deciding whether analysis of variance is appropriate in this situation, in terms of whether the underlying assumptions of the statistical model have been met. For the obtained conditions there seems to be no way available to test the assumption of homogeneity of variance from cell to cell. For example, with respect to the assumption of normality of the variate in the population, it should be noted that various plots of the action rate data indicate that our sample distributions approximate normality at least crudely. Generally, for the purposes of this study the assumptions were considered to be met well enough to warrant the use of the analysis of variance technique. The indices derived from this procedure were judged satisfactory and useful for showing general trends and avoiding gross errors, if the probability figures assigned to them are not taken too literally.

Table 16

Inter-problem Consistency Coefficients for
Action Rates of Subjects in Homogeneous Format-Gear Groups

Data Group	Number of Subjects	No. of Identical Problems on Which Tested	Inter-problem Consistency Coefficients	
			All Possible Pairs of Problems	All Problems Combined
Job Sample Radio	24	6	.47	.84
Job Sample Radar	24	6	.49	.85
MASTS Radio	24	6	.31	.73
MASTS Radar	24	6	.51	.86
AUTOMASTS Radio	45	5	.46	.81
AUTOMASTS Radar	45	5	.63	.90

problems, as computed by the intraclass correlation formula. These coefficients range from .31 to .63. If we consider each subject's action rate on a problem as his score on that problem, then his average rate score for the combined set of k problems would tend to be more stable. The coefficients in the second column range from .73 to .90, and are equivalent to the reliability coefficients which would be obtained if each subject's rate scores on the set of k problems were combined into a composite rate score. In this sense the coefficients are equivalent to the results that would be obtained by applying the Spearman-Brown formula to an intraclass r to predict reliability for a measure k times as long (12).

Since the above correlations are based on groups tested on the same format and gear, the consistency in rates is not related to either of these factors. At the least, then, if we consider only problems within the same format and gear, subjects show moderate consistency in working speed from problem to problem. To this extent we are justified in speaking of "fast" or "slow" trouble shooters. Next, let us examine the test problems used in these homogeneous format-gear situations.

<u>Action Rate Differences Between Problems</u>

To test possible problem differences in rate within the homogeneous subgroups, six F tests were carried out, using the same analysis of variance tables previously involved in computing the consistency coefficients. Table 17 on page 144 summarizes the results.

We see from Table 17 that rates of action differ significantly between problems in each group, at the .05 level of significance. The actual differences in average rates between problems were generally quite small, of course. From all six data groups, out of a total of 34 problems, the average problem rates ranged only from 1.84 to 4.15 actions per minute. According to the F tests, however, these differences were large enough and stable enough for us to conclude that rates are significantly affected by these problem differences.

Because of our reservations about the applicability of analysis of variance techniques, an additional test of the rate differences for problems was made, utilizing a non-parametric method.

Table 17

F Tests of the Differences in Action Rates Between Problems
(Based on data summarized in Table I of Appendix C)

Data Group	F
Job Sample Radio	9.88**
Job Sample Radar	4.65**
MASTS Radio	2.86*
MASTS Radar	9.37**
AUTOMASTS Radio	2.57*
AUTOMASTS Radar	17.42**

* Significant at .05 level
**Significant at .01 level

In this case the only assumption underlying the statistical test is that the individuals can be considered as randomly drawn from a parent population, whatever the form of this population distribution may be. For our data this assumption is tenable, and the subsequent significance tests are more strictly applicable than are conventional F test methods. Mood's distribution-free test of correlated sets (35) was applied to the rate data for the six aforementioned groups, with the results shown in Table 18 on page 145.

With the exception of the AUTOMASTS Radio problems, the obtained chi square values indicate statistically significant differences among problem rates for each format-gear group.

Table 18

Chi Square Tests of Inter-problem Differences in Action Rates
(Based on same data as in Table 17)

Data Group	Degrees of Freedom	Chi Square
Job Sample Radio	5	22.08***
Job Sample Radar	5	14.72*
MASTS Radio	5	11.53*
MASTS Radar	5	36.53***
AUTOMASTS Radio	4	7.41
AUTOMASTS Radar	4	44.30***

*   Significant at .05 level
***Significant at .001 level

Table 18 agrees rather closely with Table 17. It seems clear that the magnitude of any given action rate is dependent on the problem from which it was obtained.

<u>Action Rate and Problem Success</u>

Differences in trouble shooting speed between successful and unsuccessful performances have been mentioned in several respects so far in our discussion. Such rate differences were first noted in the summary table on page 138 showing action rates for all major subgroups. Successful and unsuccessful performances differ markedly as to action rate when the averages for any given problem are considered. Taking each problem separately by format, in 45 out of 50 possible instances the average rate for successful performances

is higher than for unsuccessful performances on the same problem. A chi square test of these results, based on a sign test of correlated scores (35), yields a value of 32.00 for 1 degree of freedom, a result which could occur by chance alone less than one time in a thousand.

A somewhat different question is: On the average does <u>each subject</u> work at a faster rate on problems which he solves than on those he fails? As in the preceding analysis, a sign test was applied. Results are shown in Table 19 below.

From this table we conclude that Job Sample and MASTS subjects work faster on problems which they solve than on those they fail. This result is not demonstrated for the AUTOMASTS format.

Table 19

Sign Tests of Action Rate Differences Between Each Subject's Successful and Unsuccessful Performances

Data Group	Chi Square (Based on 1 degree of freedom)
Job Sample Radio	15.38***
Job Sample Radar	7.53**
MASTS Radio	25.00***
MASTS Radar	24.03***
AUTOMASTS Radio	1.88
AUTOMASTS Radar	0.23

** Significant at .01 level
***Significant at .001 level

Radio-Radar Differences in Action Rate

Another question is whether there are rate differences between a trouble shooter's radio and radar performances. The time analyses so far have generally presented separate rates for the different equipments. This separation was maintained because of the differential average rates shown in the summary table.

A sign test of this relationship was made by pairing each individual's average radio rate against his average radar rate. The chi square obtained from this test was 20.75, statistically significant at the .001 level for 1 degree of freedom.[86]

In this case where subjects are all working on identical problems, it is seen that each man performs on the radio problems at significantly higher average rates. The average difficulty of the radio problems was about the same as the average difficulty level for the radar problems. They were matched in other ways also. Thus, it appears that the difference in rates is directly due to the types of gear being worked on. Perhaps these men were more familiar with radio circuitry and could formulate their checking procedures more readily.

Action Rate and Problem Order

Do subjects tend to get faster or slower as the problem series proceeds? Since both Job Sample and MASTS used counterbalanced

---

[86] In this calculation average rates were computed for all 81 subjects on the three formats, using the same five radio and five radar problems that were used on the AUTOMASTS. This was done so as not to confound the average rates with problem differences, but resulted in the mixing of performances on Job Sample and MASTS formats for 36 individuals. Because format effects were partially counterbalanced in this arrangement (21), the ambiguity so introduced was considered less than would occur if rates were based on different problems for each subject.

designs, these two data groups were used in this analysis. Plots of the average rates for each sequence position, one through six, suggested a trend of slight increases in average rate from one problem to the next in each format-gear situation. However, chi square evaluations by Mood's test of correlated sets showed no significant difference between sequence positions for any of the four format-gear situations. The order in which a problem was taken in a series apparently was not a significant determiner of working speed.

<u>Action Rate as a Predictor of Trouble Shooting Goodness</u>

The discussion to this point has sought to characterize and predict the overall action rate from knowledge of other factors. In many practical situations we also would like to be able to reverse the prediction. That is, we would like to predict other variables from knowledge of average action rate. Of particular interest are predictions involving the general question, "Is overall speed an indicator of trouble shooting goodness?"

This general question can be broken down into subquestions. For example, does a man who goes fast make more errors than a man who goes slow? This was investigated with respect to the Job Sample and the AUTOMASTS formats. The results are reported in Table 20 on page 149.

On the basis of the correlations presented in Table 20 it appears that there is little relationship between average action rate and average number of errors on the Job Sample test. A slow technician working on the electronic equipment was as likely

Table 20

Correlations Between a Man's Average Number of Errors and His Average Action Rate

Format	r Minor Errors	r Major Errors
Job Sample (N = 36 ETs)	-.02	-.20
AUTOMASTS (N = 45 ETs)	.61	--

Note: Each man's averages are based on 12 Job Sample or 10 AUTOMASTS problems.

to make errors (both major and minor) as a man with a faster average action rate. However, on the AUTOMASTS, speed and errors (on the AUTOMASTS only minor errors could be made) were significantly related. Furthermore, for this same AUTOMASTS group a significant correlation of .42 was obtained between a man's average action rate and the average number of errors made per minute that he worked. Thus, it appears that when a man works faster he tends to make more errors, so that the old adage "haste makes waste" is true in the AUTOMASTS situation.

A second issue involves the relationship between a man's average action rate and the Direct Clue Actions he performs. As described in our Technical Report No. 12, Direct Clue Actions are checks which, if properly interpreted, definitely localize the trouble to a small area of the equipment, usually to one or two components. The Direct Clue Actions score was defined as the

ratio of the number of Direct Clue Actions to the total number of actions. This score was shown to be rather highly related to trouble shooting goodness as defined in several different ways. For the AUTOMASTS data, it is negatively related ($r = -.41$) to action rate. Faster AUTOMASTS rates thus are less efficient than slower rates in terms of the number of Direct Clues per action.

Direct Clue Actions can be treated another way by dividing the total number of Direct Clue Actions a man makes in all ten performances by the total number of minutes that he works. This revised index represents essentially the same measure of trouble shooting progress as the Direct Clue Actions score, but changes the base to one which is more directly comparable with actions per minute. For the 45 technicians tested on 10 AUTOMASTS problems, the correlation obtained between each subject's average rate and Direct Clue Actions per minute is .28. From this insignificant coefficient we conclude that rate of action by itself is not useful in predicting whether a man will find many Direct Clues per unit of time.

The most frequently used indicator of "goodness" on trouble shooting tests is the proportion of problems correctly completed. Correlations between the proportion of problems solved and average action rate for each ET are reported in Table 21 on page 151.

Coefficients in Table 21 are based on the maximum number of subjects tested on identical problems in each of the three formats. Correlations are shown between a man's proportion of successes and each of three average action rates which were calculated for

Table 21

Correlations Between a Man's Average Action Rate and the Proportion of Problems He Solves

Format	Number of ETs	Number of Problems	Action Rates Based On All Problems	Solved Problems	Unsolved Problems
Job Sample	24	12	.27	.09	.26
MASTS	24	12	.28	.09	.14
AUTOMASTS	45	10	.19	.18	.10

him: his average rate on all of his problems, on those he solved only, and on those he failed only. None of these correlations is statistically significant. This indicates that a man's success in solving trouble shooting problems is not closely related to, or predictable from, his average working speed.[87]

This conclusion is supported by correlations between shipboard ratings of trouble shooting proficiency and average action rates obtained on the Job Sample and MASTS tests. The proficiency

[87]Strict comparisons using those AUTOMASTS subjects who failed C115 and those who solved C111 provided a means for comparing the average rates of the AUTOMASTS solvers and non-solvers with no possibility of the results being affected by variable problem differences. Each trouble shooter was assigned to an "above median" or "below median" group on the basis of the proportion of nine problems that he solved. These two groups were then compared as to whether their average rates were above or below the median of the unsuccessful rates on problem C115. A 2x2 chi square test indicated that the differences between the two groups could have occurred by chance alone. A parallel comparison based on an easy problem (C111) yielded essentially the same results.

ratings were made by the electronics material officers who had direct shipboard responsibility for the men. For 30 subjects the correlation between each man's average action rate and his shipboard rating was .25 when based on Job Sample performances only, and -.03 when based on MASTS performances only. Neither coefficient is statistically significant.

The general finding is that overall working rate does not appear to be a useful predictor of trouble shooting goodness; a technician's working speed is no guide to his competence. In evaluating this result it should be noted that a man's working rate during a performance changes continually in response to the problem situation. An overall measure which combines all these changes into a single index results in considerable loss of information so that good trouble shooting attempts cannot be reliably distinguished from ineffective ones. As the following indicates, time and rate are better understood when qualitative characteristics within the performance are taken into account.

Time Relationships Within Performances

The time analyses presented so far could have been made without any knowledge of the trouble shooter's behavior except the number of actions in, and the duration of, each performance. Although the overall rate of action is a convenient index of a man's average working speed, it tells us nothing about the way that his actions are distributed in time.

Since each Job Sample and MASTS performance record contains a complete account of minute-by-minute trouble shooting activities, they provide a good source of data from which to study tempo changes

within a performance.[88] Either format could have been used in the tempo analysis, but the MASTS was selected for several reasons. Foremost among these was the fact that the MASTS data are more standardized in the sense that the possibilities for action were more restricted. It was felt that extreme differences in trends between formats would not appear. Preliminary plots of Job Sample records resembled corresponding MASTS data with respect to discernible tempo trends, thus lending support to the use of a single format for the purposes of this analysis.

Once the format was selected, a decision had to be made regarding the problems to be studied. Nine of the 30 available MASTS problems were common to all formats. Consequently, this group of four radio and five radar problems was chosen. Three of the radar problems had been taken by a group of 12 experienced electronics technicians. The remaining problems were taken by another group of 24 ETs.

The MASTS time limit was 20 minutes, so each performance record was transcribed to a plot of activities on a standard 20 minute time base. The plot showed by code numbers the actions initiated during each minute of each performance. These individual tempo graphs served as the basic working data for most of the time analyses presented below.

---

[88] However, the crudity of the time recording procedures employed leaves much to be desired with respect to ideal time data. The present effort is intended to explore, in a preliminary way, the potentialities of time analyses. On the basis of this experience, more precise studies of the time variable have been planned.

Average Intra-performance Tempo

From a cursory examination of individual tempo graphs it is evident that subjects vary considerably from one minute to the next with respect to the number of actions they perform. It is quite rare to see a plot which proceeds at a steady paco. Similarly, and not unexpectedly, subjects behave quite differently, one from another, in the temporal patterning of their performances.

An individual's tempo pattern for a given performance consisted of a minute-by-minute graph with "time" on the horizontal axis and "number of actions" on the vertical axis. A number of attempts were made to discover general trends among these individual graphs by inspection and by superimposing a number of individual tempo curves. Most curves were so variable that smoothed graphs were employed. The smoothing was accomplished by means of a three-minute moving average. The process for smoothing the individual curves is as follows: (a) Mark off successive minutes on a 20-minute horizontal base. (b) Establish a "number of actions" vertical axis. (c) Plot the number of actions occurring in the first <u>three</u> minutes as the value for the <u>first</u> minute. (d) Continue in this manner to plot points for each minute, always recording the number of actions occurring in the minute in question, <u>plus</u> the number of actions in the two succeeding minutes. Curves plotted according to the above procedure are more amenable to trend analysis, and eliminate artifacts that arise from the use of the minute-by-minute recording system employed during data collection.

The smoothed tempo curves then were combined to obtain group curves. These group curves simply represented the average value

for each of the 20 minutes of the scale, where the average was based upon the non-zero entries of the individual curves. Thus each value on the group curve represented an average of all of the ongoing performances depicted by the smoothed curves. A group curve was always terminated when fewer than three smoothed curves could contribute to it.

Group curves were plotted for each of the problems for all subjects combined, for the solvers only, and for the non-solvers only. Since the group tempo curves for each problem showed only minor differences, composite tempo graphs were constructed for the radio problems as a group and for the radar problems as a group. These are presented in Figures 28 and 29 on page 156.

Inspection of these figures reveals several trends. One is that the general tempo patterns for the radio performances are very similar to the tempo patterns for the radar performances. Average action rates start at a high point and drop somewhat sharply in the early phases. Probably this is another manifestation of the fact that a man's starting technique is relatively well structured and routinized. Unsuccessful performances tend to proceed at intermediate speeds during the middle of the allotted time and to drop to their lowest point as the time limit is approached. Successful performances follow the same general pattern, but fluctuate more in average rate.[89]

---

[89]The greater variability in average rate for the solvers may be due to the fact that these averages are based on a diminishing group of subjects, since a performance is dropped from the calculation of average rates whenever a solution occurs. This not only leads to smaller N's as the minutes proceed, but also leads to a slight distortion due to the inaccurate estimate of the final "minute."

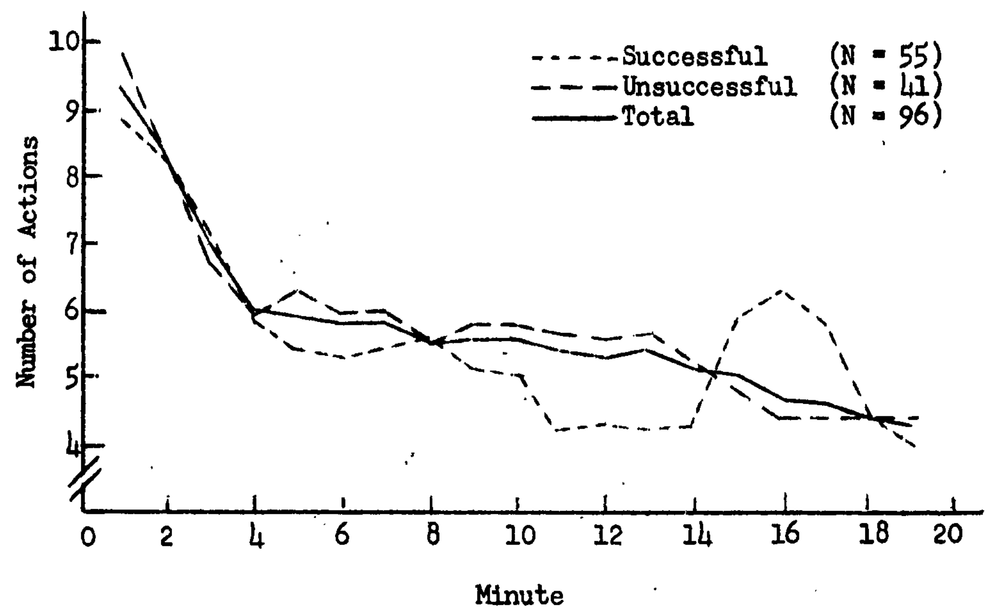


Fig. 28 Group Curves of Average Tempo on Four MASTS Radio Problems

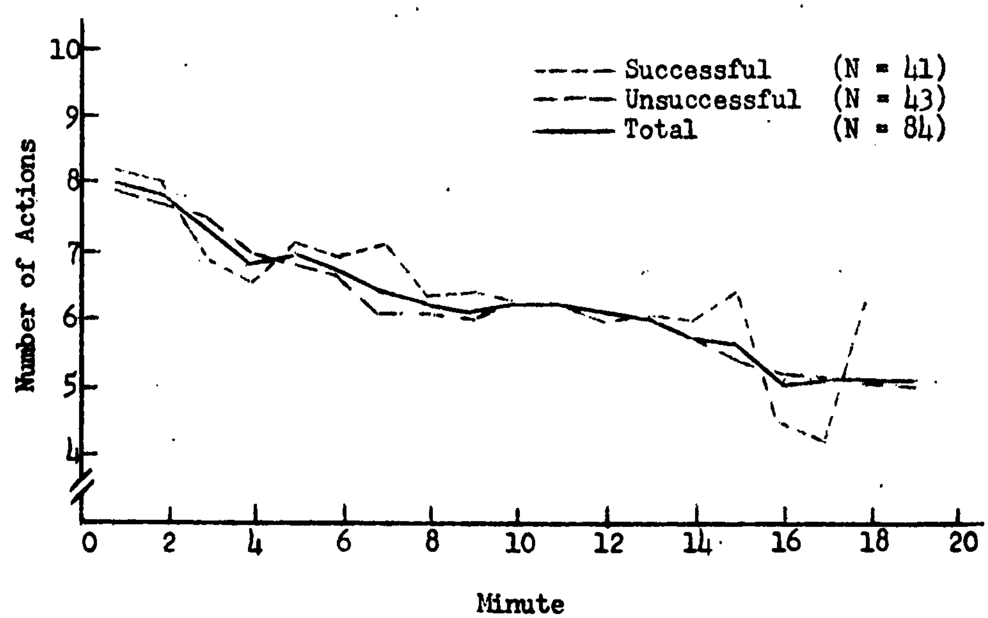


Fig. 29 Group Curves of Average Tempo on Five MASTS Radar Problems

For any single problem these trends are not so distinct. Generally the composite curve for a single problem is more irregular, and the successful and the unsuccessful curves typically intertwine. Individual tempo curves are even less interpretable. However, the group curves illustrated above do indicate, in a general sense, certain action rate trends within performances.

<u>Group Size in Successive Minutes</u>

Reference already has been made to attrition of the groups as a function of time. Figure 30 shows this relationship graphically for both radio and radar problems.

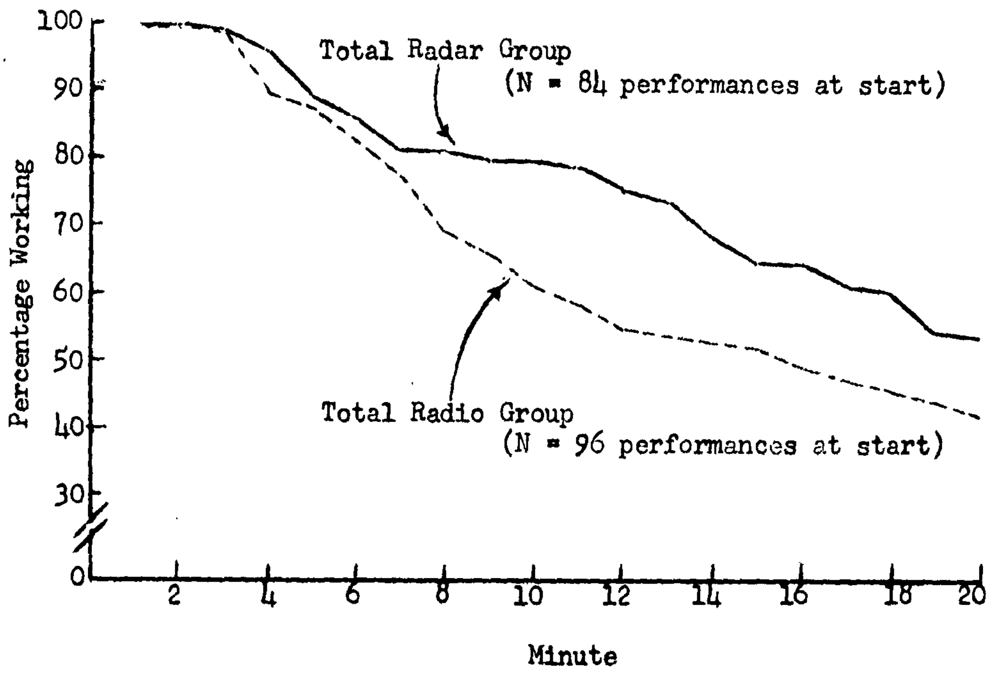


Fig. 30. Percentage of Performances Continuing in Successive Minutes. Data are from four MASTS radio and five MASTS radar problems.

Since the main reason subjects drop out is that they solve a problem, we also can see from the above curves how solutions are distributed through time. Radio and radar curves follow each other quite closely for the first seven minutes or so, and then the radio curve continues to drop while the radar curve remains about level from minute seven through minute eleven. Then the curves become approximately parallel again.

Types of Checking Activity as a Function of Time

Preceding graphs (Figs. 28 and 29) of actions per minute give a rough idea of tempo trends within a performance but they do not show the times for different phases or the way that different kinds of activities are distributed in time. Figures 31 and 32 on page 159 present average times to the end of the initial localizing sequence (ILS), first isolating sequence (IS), and the first replacement.[90] They also give minute-by-minute breakdowns for the relative popularity of each type of action.

The average ILS for the problems included in Figures 31 and 32 takes about three minutes. On the basis of the overall times spent on these problems, the ILS extends about a fourth of the way through the performance. About a minute later the technician finishes his first IS, and by the end of the sixth minute he makes

[90] It will be recalled that the ILS includes all actions up to first entry into a stage, and that an IS consists primarily of voltage and resistance checks within a single stage. In calculating times to the ends of these phases, it frequently happens that no time entry is available for the exact action where the phase ends. It is therefore necessary to estimate fractions of a minute. Such estimates are based on the number of actions occurring in the minute; barring information to the contrary, it is assumed that the actions in a given minute are evenly distributed in time.

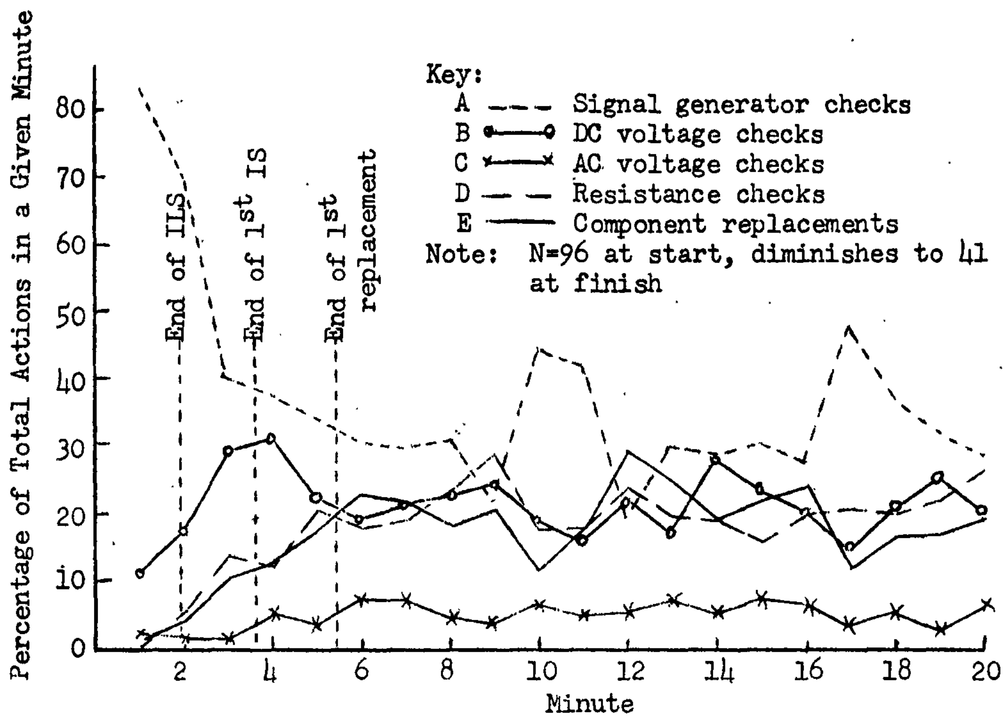


Fig. 31 Time Plots of Radio Trouble Shooting Activities Classified by Type of Action. Data are from four selected MASTS problems.

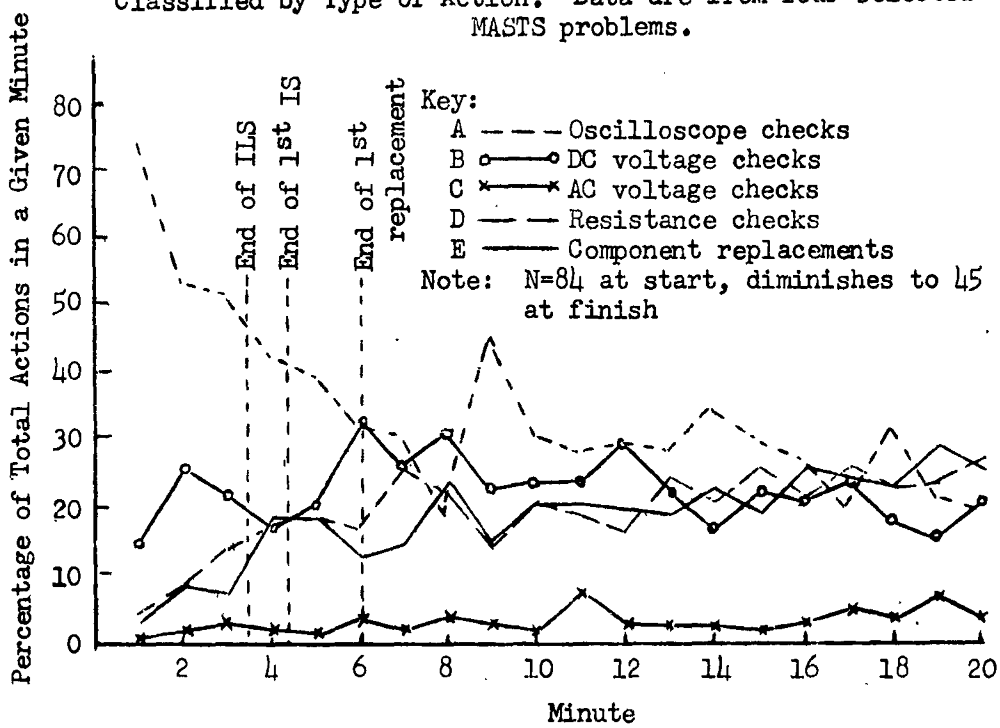


Fig. 32 Time Plots of Radar Trouble Shooting Activities Classified by Type of Action. Data are from five selected MASTS problems.

his first replacement. These times are, of course, parallel to the average number of actions (7.6, 10, and 14, respectively) required to reach these phases. Also, the two-minute span between completion of the first IS and the first replacement is related to the fact that most technicians do not proceed immediately from isolating checks to a replacement.

We have already seen that the men work fastest at the beginning of a problem. The ILS series of localizing signal and waveform checks, then, must be the fastest phase. Does this hold up throughout—are localizing efforts extending across a series of stages faster than checks inside a stage after localization has been achieved? When all localizing sequences are compared with all within-stage check series, the rates are very nearly the same, and subsequent statistical tests showed no significant differences. On this evidence, we cannot support the notion that a technician localizes to a stage at a relatively high rate and then slows to a more deliberate pace inside the trouble area.

What about the types of activity made in successive minutes? Figures 31 and 32 were compiled by taking all the actions made in each minute and then plotting them according to the proportion of each type of action. The only kinds of checks which vary much are signal and waveform readings. In the first minute, they account for about three-quarters of the activity; this proportion drops steadily through the ILS, first IS, and first replacement. A couple of minutes after the first replacement, signal and waveform checking takes a decided jump in popularity, with one or more little bursts later on. These later accelerations, if they are significant,

probably reflect a group tendency to set up a new localization sequence after an unsuccessful series of actions within a stage. DC voltage, resistance, and replacement behavior stabilizes in popularity after the first two or three minutes. It was expected that the replacement curve might show a relative upturn near the time limit, but this did not happen--there is no particular tendency for clusters of long-shot guesses to occur in the last few minutes.

Spurts and Lags

As a technician works along, he seldom maintains a steady action rate. Sometimes he checks in a rapid sequence, sometimes he pauses momentarily, or proceeds in a halting fashion. If sufficiently sensitive (accurate to ± 1 second) time data were available, one would expect to be able to relate these periods of differential activity to qualitative characteristics of the performance. Among other things, such data should indicate where the man was "stumped," where he was uncertain, or where he was relatively sure of himself. Furthermore, accurate timing should shed additional light on the composition of the "behavioral units" referred to in Section III.

While the present data are crude (accurate to ± 30 seconds) a preliminary examination was made on a sample of 180 MASTS records. Arbitrary criteria of spurts (any minute with seven or more actions) and lags (a three-minute actionless interval) were set up. These hyper- and hypo-active periods were identified for each performance. Inspection of these records leads to the following observations:

a. Most spurts appear near the beginning of a trouble shooting performance. The typical spurt was about three-tenths of the way through the performance, with about half of them occurring within the first two minutes.

b. Lags are generally much later, with the typical lag occurring after three-quarters of the performance has been completed. In about a third of the cases the lag extends to the very end, showing that the ET was completely stumped.

c. Most often the actions in a spurt are not confined to a single stage. They usually consist of signal generator and oscilloscope checks along a signal or pulse path.

d. There seem to be two types of lags: "quitting" lags and "thinking" lags. During a "quitting" lag the technician appears to have abandoned any serious attempt to solve the problem. This type of lag occurs mainly at the end of a performance and appears to be simply a case where the man is waiting for the time to run out. The "thinking" lag is also late in a performance, but seems to be devoted to reviewing previous activity and studying reference materials. Although the technician is not making any checking actions he still appears to be busily engaged in trouble shooting.

e. In most cases, the checks a man makes two minutes before a lag are distributed through several equipment stages, but in the two minutes after a lag only one stage is checked. Component replacements are likely following a lag.

SECTION VI. SUMMARY AND CONCLUSIONS

The purpose of this section is to present a summary of the research and a list of the principal conclusions.

Summary

Objectives

Three general objectives of the report may be distinguished. The first is to present a treatment wherein primary attention is focused on the full sequential records of trouble shooting. This

implies that serious efforts be made to deal with the complexities of such data, to define critical aspects of the behavior, and to determine the consequences of different action patterns. It is expressly recognized that such an analysis be carried forward without primary reference to test or scoring applications, since those aspects have been examined previously.

A second objective is to integrate trouble shooting information and analytic approaches not previously reported. Some of the data and methods discussed here have been accumulating for several months. It is felt that they deserve publication, especially in view of the current research interest in corrective maintenance problems.

The third objective is to make, insofar as possible, general statements about trouble shooting. The fact that different problems and formats are represented in the data affords an opportunity to find out which conclusions hold up throughout and which ones are closely bound to special conditions. Such comparative data have been very scarce in the trouble shooting field. Even within the present limitations, it is believed that such an orientation provides a modest test of the generality of the findings.

Procedures

The analytic procedures are contained in three main chapters, as shown in the following outline:

1. Progressive phases of trouble shooting:

   a. Separation into phases. Each performance was broken down into qualitative segments, as defined by a schema. A segment continued until some activity was encountered which met the requirements for another type of segment (e.g., an Initial Localizing Sequence continued until localization to a stage was reached.).

b. Intra-segment analysis. The internal characteristics of each segment were examined (e.g., length, frequency of occurrence, placement in the performance, individual consistency across a series of problems, etc.).

c. Inter-segment analysis. The characteristics of each class of segment were related to other kinds of segments (e.g., the average length of a stage-localizing sequency was compared with the amount of subsequent within-stage activity).

d. Effects. The consequences of the behavior in each segment were examined, insofar as the total performance was concerned (e.g., relationships between initial search pattern and likelihood of success).

2. Intensive treatment of specific trouble shooting issues:

a. Utilization of information. Each performance was traced through and those points where certain information was achieved were labeled. The consequences of obtaining, or not obtaining, these checks were investigated.

b. Rate of convergence. A "distance from the trouble" weight was applied to each action in each performance. Changes in distance were plotted and a least-squares line was fitted. The slope of the line was taken as a "rate of convergence" index and was correlated with other performance measures.

c. Predicting expert judgment. The trouble shooting records were ordered according to the "trouble shooting quality" ratings assigned by electronics experts. The single simple performance index which best predicted the experts' ordering was found. Other variables were tried out until a system was developed which reproduced the original ordering by successive sorting on a few simple performance characteristics.

d. Redundancy. Eight classes of redundant trouble shooting behavior were identified. The occurrence and variability of these classes in the behavior records were observed. The behavioral significance of literal repetitions and virtual repetitions was compared.

e. Errors. Each record was searched for errors. Major errors (e.g., blown fuse due to carelessness in using test equipment) and minor errors (e.g., injection of signals into radio power supply) were identified. Subsequent effects of making these kinds of errors were investigated.

f. Trouble shooting and problem solving. The literature on problem solving behavior was reviewed, and parallels between laboratory problem solving research and the trouble shooting situation were discussed.

3. Temporal factors in trouble shooting:

a. Standardization. All records were adapted to the same system of description (e.g., Job Sample activities such as "feels tubes" or "looks for tube manual" were eliminated). Minute-by-minute plots were made.

b. Determination of times and rates. The times required to reach qualitative phases were tabulated to the nearest minute (e.g., first replacement, end of performance, etc.). Actions per minute was taken as the overall rate base.

c. Specific rate hypotheses. The significance of rate and time was tested in terms of such factors as problem conditions, equipment difference, success-vs.-failure, individual consistency, series effects, etc.

d. Intra-performance trends. Group rates were plotted in successive minutes by means of the moving average method. The differences in average rate trends between successful and failing performances were investigated. Relative popularity of different types of activity for successive minutes was plotted.

e. Spurts and lags. Segments of records which showed extremely high or low action rates were selected. The conditions and consequences of each kind of extreme rate were identified.

Conclusions

Sixty-two conclusions are listed below. These statements represent an effort on the part of the investigators to integrate the varied information presented in earlier portions of the report. For the reader's convenience, related statements are grouped together. These groupings and the order of the conclusions have no particular significance.

Page references accompany most of the statements so that the reader may turn directly to a more complete discussion of the point and determine more precisely the basis for the conclusion. Care should be exercised to avoid unwarranted generalization. In many cases specific limitations are pointed out in the cross reference.

IN GENERAL.....

1. Experienced technicians show marked individual differences in their ability to locate defects in malfunctioning electronics equipment.

2. The actions in a trouble shooting performance are seldom random, but are dependent on the circuitry, the problem conditions, and the subject's style of search.

3. The typical trouble shooting attempt is made up of three qualitatively different kinds of behavior: generalized searching, localized searching, and component adjustment or replacement.

4. If a man is a good radio trouble shooter, the chances are that he is a good radar trouble shooter, granting some previous exposure to each type of equipment.

GETTING STARTED.....

Approximately the first third of the average trouble shooting performance is devoted to generalized localizing activity. In the present framework this extends from the first action to the point where the man begins intensive isolating checks within a stage. Most of the activity consists of signal injections and waveform checks, and several stages are usually involved.

5. ETs begin most of their trouble shooting attempts as if they understand the general nature of the original symptoms and the basic characteristics of the malfunctioning equipment. (p. 12)

6. Most radio trouble shooting performances can be classified according to four systematic starting procedures. (p. 19)

7. The way a man starts a radio problem is not directly related to the specific trouble symptoms or other problem conditions. (p. 25)

8. The way a man gets started on one radio trouble shooting problem is predictive of the way he will start other radio problems. (p. 21)

9. On the average, those radio trouble shooting performances which start out with the "Half-split" method of attack are shorter than other radio performances. (p. 27)

10. The method a man uses at the beginning of his radio problems is not predictive of the method he will use at the beginning of his radar problems. (p. 37)

11. Most radar trouble shooting performances can be classified according to five systematic starting procedures. (p. 30)

12. In radar performances the pattern of actions the technician chooses is influenced by the particular symptoms and conditions of each problem. (p. 34)

13. A man's starting procedure on one radar problem is not predictive of his starting procedure on his other radar problems. (p. 32)

14. The way a man starts a trouble shooting problem is not predictive of whether or not he will find the trouble. (p. 24)

15. Men independently tend to alter their starting behavior in a desirable direction as a result of taking a series of trouble shooting problems. (p. 17)

16. The way that a man starts to work on a problem is independent of his previous success or failure with that particular starting method. (p. 24)

17. In both radio and radar problems, the most popular method of approach is to start at the back (output) end of the equipment and to go systematically toward the front (input) end of the equipment, along the signal or pulse path. (pp. 20,31)

18. If a man makes relatively few checks before concentrating his activity to a single stage of the equipment on one problem, he will tend to make relatively few checks of this type on his other problems. Similarly, a man who makes many checks previous to entering a stage on one problem will make many such checks on his other problems. (p. 42)

19. There is no particular advantage in accomplishing a large number of checks before restricting checking activities to a single equipment stage. (p. 42)

20. In a large proportion of the performances enough information is gained to indicate clearly the malfunctioning stage. (p. 83)

21. Acquisition of sufficient information to locate the defective stage does not guarantee a correct solution. However, it does increase the chances of success. (p. 83)

GETTING CLOSER.....

Localized searching behavior within a restricted area of the equipment takes up about a third of the average trouble shooting attempt. It usually reflects the technician's belief that his generalized search has narrowed the trouble area. The typical performance contains three sequences of intensive checking within a stage. These intra-stage sequences are short and consist mainly of DC voltage and resistance measurements. Generally, two such sequences occur before the first replacement is made.

22. As a performance progresses, there is a tendency for the technician to get closer to the trouble, whether or not he ever finds it. (p. 91)

23. In a large proportion of the performances, enough information is gained to make, theoretically, positive identification of the defective component. (p. 88)

24. The number of times that a man makes consecutive within-stage checks on one problem is predictive of the number of such sequences he will perform on other problems. (p. 45)

25. The length of within-stage checking sequences does not progressively change as a performance proceeds. (p. 46)

26. Checking sequences within the confines of a single equipment stage are relatively unsystematic. (p. 46)

27. There are no standard patterns from which one can predict when a man will make a component replacement. (p. 69)

28. A tendency to make more checks in "wringing out" a stage improves a man's chances of locating the trouble. (p. 45)

29. The number of checks a man makes before he replaces a component is relatively independent of the type of equipment, test format, output symptoms, or the order in which the problems are administered. (p. 48)

30. The number of actions an ET makes before his first replacement on one problem is predictive of the number of pre-first-replacement checks on his other problems. (p. 52)

31. There is no particular advantage in making a large number of checks before making the first replacement. (p. 52)

32. On harder problems, technicians make more checks before they perform their first component replacement. (p. 48)

33. If a technician switches from the use of one type of test equipment to another, for example, from signal generator to ohmmeter, it is unlikely that he will replace a component within the next three actions. (p. 70)

THE PAYOFF.....

Nearly every performance features at least one component replacement; the average ranges from two to four. Most initial replacements take place after a series of generalized and localized search activities--at about the middle of the performance, in the typical case. In a sense, a replacement represents an integration of a man's previous searching behavior and serves as a check on his interpretation of the problem data.

34. Most initial replacements are "plausible", in the sense of being either the defective component or a reasonable choice on the basis of the available problem information. (p. 60)

35. The "nearer" a man's first component replacement is to the defective component, the more likely it is that he will eventually find the trouble. (p. 60)

36. The relative frequency with which different types of components are replaced does not correspond to the frequency with which the parts are represented in the equipment. (p. 54)

37. There is a pronounced tendency to replace tubes. (p. 54)

38. If a man makes a replacement which does not correct the trouble symptoms, his next few actions are usually close to the component he has just replaced. (p. 71)

39. In general, a component replacement is not followed immediately by another replacement. (p. 67)

40. Apparently the post-replacement reaction is a "spur of the moment" matter, and is not clearly related to personal styles of trouble shooting. (p. 72)

41. Component replacements that occur consecutively are usually located close together in the equipment. (p. 68)

42. The likelihood of success on a problem diminishes with successive replacements, so that after a man makes two or three replacements, it is very unlikely that he will find the trouble. (p. 66)

43. There is no indication that technicians resort to wild or implausible replacements as the performances proceed. (p. 65)

REDUNDANCY.....

A redundant action is one which furnishes no new information--one which if omitted would leave the performance essentially complete. Practically every trouble shooting attempt contains some redundant activity. The proportion of redundant actions may range as high as 75 per cent, and the average is 30 to 50 per cent of the performance.

44. Two qualitatively different types of redundant behavior are distinguishable. (p. 98)

45. If an ET has a high proportion of redundancies on one problem, he will have a high proportion on the rest. (p. 103)

46. When all performances are transformed to a standard number of actions, each type of redundancy has its own distinctive distribution curve. (p. 106)

47. Literal repeats are much more frequent than virtual repeats, when men are working on actual electronics equipment. (p. 101)

48. In trouble shooting on symbolic formats there are about as many literal repeats as virtual repeats. (p. 101)

49. Many of the virtual repeats appear to be due to incorrect interpretations of signal flow relationships. (p. 104)

50. There is a slight negative relationship between the proportion of redundant behavior in a performance and conventional criteria of performance goodness. (p. 114)

TEMPO.....

Time limits for the problems used here vary from ten to thirty-five minutes; the time spent on a typical problem is about fourteen minutes. A man's working rate changes continually in response to the problem situation. The average rate is two or three actions per minute.

51. Technicians are moderately consistent in their working speed from one problem to another. (p. 142)

52. Rate of checking is typically faster at the beginning of a performance than at any other time. (p. 155)

53. Extremely low action rates are most likely to occur near the end of a performance. (p. 162)

54. Speed of checking is influenced markedly by the particular problem being worked on. (p. 143)

55. Trouble shooting on radio receivers is accomplished at a faster rate than trouble shooting on radar-type circuitry. (p. 140)

56. The order in which a problem is taken in a series is not a significant determiner of working speed. (p. 148)

57. A man's average working speed is not generally useful for predicting the number of problems he solves. (p. 152)

ALSO.....

58. Most trouble shooting performances feature some directed trial and error behavior. (p. 127)

59. Expert judgments of the goodness of a trouble shooting performance appear to depend heavily on (a) whether or not the performance terminated in solution, (b) the number of actions in the performance, and (c) the closeness with which the defective component is approached in the final minutes. (p. 96)

60. In a sizeable proportion of the repair attempts involving live equipment, technicians engage in practices which endanger themselves or their test equipment. (p. 117)

61. Performances containing activities which endanger the ET or the equipment typically end in failure. (p. 118)

62. Almost every performance contains minor errors of test instrument usage, but these errors have little influence on likelihood of solution. (p. 119)

SECTION VII. IMPLICATIONS

As the numerous references throughout the text have indicated, the present research has attempted to keep abreast of developments in a very active field. In this section our remarks are directed toward a few broad strategic issues which relate to the utilization and furtherance of trouble shooting research in general.

At the present time the study of trouble shooting is less of a bandwagon than it was a few years ago. In many ways this is good,

since the field is being cluttered with fewer studies bearing only a tangential relationship to genuine maintenance problems. In other ways, this is unfortunate, for it has led some to regard trouble shooting research as passé. Insofar as this attitude discourages continued support of research in this area, the attitude is bad. The problems here are old, and they are difficult; but they will yield to persistent effort. Questions associated with directed thought and judgment are among the most intractable in psychology. The fundamental units and categories for coding the behavior need to be systematized and elaborated further, so that critical features can be pinpointed.

But these facts should not obscure the progress that has been made. It has not been spectacular. It has not eliminated all bad personnel practices or bad equipment design. It has produced a cadre of experienced trouble shooting researchers and a number of devices and techniques which have proved to be useful. Particular gains have been made on evaluating and scoring trouble shooting proficiency. It is now possible to assign scores to trouble shooting performances with considerable precision and confidence. Also, there is a growing body of specific information regarding the ways that performance characteristics relate to each other. In other words, there is an accumulating fund of knowledge available for continued work in this domain.

A major difficulty is that each study is so specific that it is very difficult to integrate all the findings. This is a result of differing objectives, varying contexts, practical restrictions, and the fact that everybody started from scratch at about the same

time. In many ways these differences are beneficial, and it would be premature to advocate a doctrinal approach. Nonetheless, some way must be found to organize the assorted results. One approach is to subsume trouble shooting under an established topic, such as learning, and to derive specific hypotheses and general laws from the learning model. Efforts have been made in this direction, and it is well that they be continued. However, it may be unwise to put all of the trouble shooting eggs in the learning basket. Integration on other fronts should be attempted. In particular, the possibility for long-range cooperative efforts between investigators should be explored.

Finally, it is of utmost importance that the integration and systematization of trouble shooting knowledge be based on awareness of and continued contact with trouble shooting as it exists in the field. The researcher must not retire to an ivory tower while building behavioral theories of trouble shooting or investigating obscure points. Although trouble shooting activities do provide an excellent source of problem material for basic studies in human problem solving ability, the researcher must not lose sight of some more immediate practical objectives. He must realize that only through conscientious, cooperative research efforts can there be a sound basis for guiding personnel policy and equipment design in the electronics field.

As electronics becomes increasingly important and pervasive in its military and civilian applications--a trend which appears inevitable--the personnel problems associated with corrective maintenance will grow in scope and urgency. The need for a systematic

framework for analyzing trouble shooting will become even more compelling. It is up to the research worker to meet this challenge and to provide a solid research background for future policy decisions which will have to be made.

REFERENCES

1. Berkshire, J. R. Field evaluation of a trouble-shooting aid. San Antonio: Lackland AFB, AFPTRC-TR-54-24, March 1954.

2. Billings, M. I. Problem-solving in different fields of behavior. Amer. J. Psychol., 1934, 46, 259-292.

3. Bryan, G. L. The AUTOMASTS: an automatically-recording test of electronics trouble shooting. Los Angeles: Univer. Sou. Calif., Rep. No. 11, Contract Nonr-228(02), NR 153-093, August 1954.

4. Bryan, G. L. Certain psychometric and scoring problems in trouble shooting performance testing. Paper read at Amer. Psychol. Ass., San Francisco, September 1955.

5. Bryan, G. L. An experimental battery for measurement of the proficiency of electronics technicians. Los Angeles: Univer. Sou. Calif., Rep. No. 12, Contract Nonr-228(02), NR 153-093, March 1955.

6. Carhart, R. R. A survey of the current status of the electronics reliability problem. Santa Monica: Rand Corp., Res. Memo. No. 1131, August 1953.

7. Cornell, F. G., Damrin, D. E., Saupe, J. L., & Crowder, N. A. Proficiency of Q-24 radar mechanics: III. The tab test of trouble-shooting proficiency. San Antonio: Lackland AFB, AFPTRC-TR-54-52, November 1954.

8. Crowder, N. A. Proficiency of Q-24 radar mechanics: V. Level of trouble-shooting performance observed. San Antonio: Lackland AFB, AFPTRC-TR-54-102, December 1954.

9. Crowder, N. A., Morrison, E. J., & Demaree, R. G. Proficiency of Q-24 radar mechanics: VI. Analysis of intercorrelations of measures. San Antonio: Lackland AFB, AFPTRC-TR-54-127, December 1954.

10. Damrin, D. E. & Saupe, J. L. Proficiency of Q-24 radar mechanics: IV. An analysis of checking responses in trouble-shooting on tab test problems. San Antonio: Lackland AFB, AFPTRC-TR-54-53, November 1954.

11. Dewey, J. How we think. Boston: Heath, 1910.

12. Ebel, R. L. Estimation of the reliability of ratings. Psychometrika, 1951, 16, 407-424.

REFERENCES (Continued)

13. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.

14. Evans, R. N. & Smith, L. J. *A study of performance measures of trouble shooting ability on electronics equipment.* Urbana: Univer. Ill., Coll. Educ., Contract N6-ori-07142, October 1953.

15. Fattu, N. A. & Mech, E. V. The effect of set on performance in a "trouble shooting" situation. *J. appl. Psychol.*, 1953, 37, 214-217.

16. Fattu, N. A., Mech, E. V., & Kapos, E. Some statistical relationships between selected response dimensions and problem-solving proficiency. *Psychol. Monogr.*, 1954, 68, No. 6 (Whole No. 377).

17. Glaser, R. *A proficiency test battery for guided missile technicians.* Pittsburgh: Amer. Inst. Res., Rep. No. 3, Contract N7onr-37008, NR 152-079, August 1953.

18. Glaser, R. & Schwars, P. A. *Scoring problem-solving test items by measuring information.* Pittsburgh: Amer. Inst. Res., Memo., September 1953.

19. Glaser, R. & Phillips, J. C. *An analysis of test of proficiency for guided missile personnel. III. Patterns of trouble-shooting behavior.* Pittsburgh: Amer. Inst. Res., Tech. Bull. 55-16, Contract N7onr-37009, NR 154-079, August 1954.

20. Grings, W. W. *A methodological study of electronics trouble shooting skill: I. Rationale for and description of the multiple-alternate symbolic trouble shooting test.* Los Angeles: Univer. Sou. Calif., Rep. No. 9, Contract Nonr-228(02), NR 153-093, August 1953.

21. Grings, W. W. *A methodological study of electronics trouble shooting skill: II. Inter-comparisons of the MASTS test, a job-sample test, and ten reference tests administered to fleet ETs.* Los Angeles: Univer. Sou. Calif., Rep. No. 10, Contract Nonr-228(02), NR 153-093, August 1953.

22. Grings, W. W. *Shipboard observation of electronics personnel. A description of the research.* Los Angeles: Univer. Sou. Calif., Rep. No. 1, Contract Nonr-228(02), NR 153-093, January 1953.

REFERENCES (Continued)

23. Grings, W. W. Shipboard observation of electronics personnel: general conclusions and recommendations for further research. Los Angeles: Univer. Sou. Calif., Rep. No. 7, Contract Nonr-228(02), NR 153-093. January 1953.

24. Guetzkow, H. An analysis of the operation of set in problem-solving behavior. J. gen. Psychol., 1951, 45, 219-244.

25. Hertzka, A. F. The development and evaluation of proficiency measures in A/S sonar maintenance: II. A preliminary investigation of job-sample testing for trouble shooting. San Diego: U. S. Naval Personnel Res. Unit, PRFASD Rep. No. 63, January 1955.

26. Hill, J. H., Older, H. J., & Schaefer, W. C. An analysis of maintenance duties on jet propelled Naval aircraft. Philadelphia: Inst. Res. Hum. Relat., Contract N8onr-69406, May 1952.

27. Hill, J. H., Buckley, E. P., & Older, H. J. Post-training performances of aviation machinist mates. Philadelphia: Inst. Res. Hum. Relat., Tech. Rep. No. 5, May 1952.

28. Hoyt, C. J. Note on a simplified method of computing test reliability. Educ. psychol. Measmt, 1941, 1, 93-95.

29. Jensen, J. J., Hill, J. H., Siegel, A. I., & Courtney, J. The development of a job sample trouble-shooting examination. Philadelphia: Inst. Res. Hum. Relat., Progr. Rep., Contract N8onr-69402, January 1954.

30. Johns, E. R. Personal communication. 1955.

31. Johnson, D. M. A modern account of problem solving. Psychol. Bull., 1944, 41, 201-229.

32. Johnson, D. M. The psychology of thought and judgment. New York: Harper, 1955.

33. Lewis, D. Quantitative methods in psychology. Iowa City: Gordon, 1948.

34. Luchins, A. S. Mechanization in problem solving. Psychol. Monogr., 1942, 54, No. 248.

35. McNemar, Q. Psychological statistics. (2nd Ed.) New York: Wiley, 1955.

REFERENCES (Continued)

36. Miller, R. B., Folley, J. D., Jr., & Smith, P. R. Systematic trouble shooting and the half-split technique. San Antonio: Lackland AFB, Hum. Resources Res. Center, Tech. Rep. No. 53-21, July 1953.

37. Moses, L. E. Non-parametric statistics for psychological research. Psychol. Bull., 1952, 49, 122-143.

38. Rokeach, M. Generalized mental rigidity as a factor in ethnocentrism. J. abnorm. soc. Psychol., 1948, 43, 259-278.

39. Sargent, S. S. Thinking processes at various levels of difficulty. Arch. Psychol., 1940, No. 200.

40. Saugstad, P. Problem solving as dependent upon availability of function. Unpublished doctor's dissertation, Univer. Chicago, 1952.

41. Saupe, J. L. Trouble shooting electronic equipment. Urbana: Univer. Press, Contract 33(038)-13286, May 1954.

42. Siegel, A. I., Jensen, J. J., & Danzig, E. R. An investigation and test of the trouble-shooting ability of aviation electricians. Philadelphia: Inst. Res. Hum. Relat., Tech. Rep., Contract Nonr-1348(00), January 1955.

43. Simon, G. B. The development and tryout of a check list of observable behaviors in preflighting the B-25. San Antonio: Lackland AFB, Hum. Resources Res. Center, Tech. Rep. 53-7, April 1953.

44. Taylor, D. W. & McNemar, O. W. Problem solving and thinking. Annu. Rev. Psychol., 6, 455-482.

45. Tucker, J. A. Symposium on maintenance of electronics equipment. Washington: Advisory Panel on Personnel and Main. Res., Dept of Defense, 1955.

46. Vinacke, W. E. The psychology of thinking. New York: McGraw-Hill, 1952.

47. Wallas, G. The art of thought. New York: Harcourt Brace, 1926.

REFERENCES (Continued)

48. Weaver, W. Recent contributions to the mathematical theory of communications. In C. E. Shannon and W. Weaver, _The mathematical theory of communication._ Urbana: Univer. Press, 1949, 94-117.

49. Wertheimer, M. _Productive thinking._ New York: Harper, 1945.

APPENDIX A

Illustrations of Segmenting Techniques

Illustrations of IAS Categories 1-5

The following examples are taken from AUTOMASTS performances on radio problem C102. Schematic diagrams of the radio equipment are available in previous reports of this series (3, 21).

	Location of Actions				
Category	1st Action	2nd Action	3rd Action	4th Action	5th Action
1. Half-split	Grid of AVA	Grid of IF	Plate of IF	Input of Detector	Middle of Detector
2. Loudspeaker-to-antenna	Grid of APA	Grid of Phase Splitter	Grid of AVA	Output of Detector	Input of Detector
3. Antenna-to-loudspeaker	Grid of RF Conv.	Plate of RF Conv.	Grid of IF Amp.	Plate of IF Amp.	Middle of Detector
4. Middle-to-trouble	Grid of AVA	Input of Detector	Middle of Detector	Middle of Detector	Middle of Detector
5. Unsystematic	AC Voltage at grid of APA	AC Voltage at grid of Phase Spl.	AC Voltage at grid of Phase Spl.	DC Voltage at grid of Phase Spl.	DC Voltage at grid of Phase Spl.

Note: Except for the Unsystematic category, all checks are signal injections at the points indicated.

Illustrations of IAS Categories 6-11

These examples are taken from AUTOMASTS performances on radar problems. Stage abbreviations such as SSBO and CF refer to schematic diagrams of the radar circuit in previous reports (3, 21).

Category	1st Action	2nd Action	3rd Action	4th Action	5th Action
6. Front-to-back	Waveform at output of SSBO	Waveform at output of Step Counter	Waveform at output of TBO	Waveform at output of CF	Waveform at output of OSMV
7. Back-to-front	Waveform at input to horizontal sweep section OSC	Waveform at output of VTSG	Waveform at input of VTSG	Waveform at input of OSMV	Waveform at input of CF
8. Bracket	Waveform at input to horizontal sweep section OSC	Waveform at output of CF	Waveform at output of OSMV	Waveform on plate of OSMV	Replace condenser at 2nd grid OSMV
9. Probability	Waveform at output of VTSG	Replace by-pass condenser of VTSG	Adjust frequency of VTSG	Replace plate load resistor of VTSG	DCV at plate of VTSG
10. Single-stage	Waveform at plate of VTSG	Waveform at input to horizontal deflection plates OSC	Waveform at input to VTSG	Waveform at variable resistor in VTSG tuning circuit	Waveform at output of VTSG
11. Systemless	Waveform at plate of VTSG	Waveform at grid of SSBO	Adjust frequency of SSBO	Waveform at fuse of line voltage input PS	Adjust frequency of VTSG

Instructions for Identifying an Isolating Sequence (IS)
in a Performance Record

Starting with the first action in the record, consider each action in turn. When you find a series of consecutive actions within a single stage, bracket the series, and label it an IS if it contains one or more of the following:

(a) three or more DC voltage checks, excepting B-plus readings;

(b) three or more AC voltage checks;

(c) a combination of three or more AC and DC voltage checks;

(d) one or more resistance checks;

(e) one or more component replacements.

An IS continues until an action is made in a different stage, or until a signal injection or waveform check is made in the same stage.

APPENDIX B

A Reference Distribution

for Evaluating Consistency of Set

A Reference Distribution for Evaluating Consistency of Set

This appendix presents some of the details of a method for evaluating the consistency that performers exhibit in reacting to complex situations. The approach described here arose in connection with statistical treatment of the Initial Attack Sequence (IAS) classes described in Chapter 3, and the examples given are stated in IAS terms. It should be clear, however, that the method is completely general and can be applied to any type of qualitative data which is obtained under specifiable categorical restraints.

Let us assume that a set of mutually exclusive and exhaustive categories is available for application to certain behavior sequences. Assume further that each subject in the sample has undergone a successive series of situations--this can be, for example, a series of problems. The question we are concerned with is, does he always act in the manner defined by one of the categories or does he change categories frequently? In other words, does he exhibit consistency from one problem to another?

Let us make the question more explicit by referring to the diagram below. Here we have a situation where each subject took five radio trouble shooting problems. In each performance, a part of the starting behavior is identified as either an A type, a B type, a C, D, or E. Notice that Subject 21 has a frequency of five in Category C. This means all his performances began with a C-type approach. Subject 22 splits his preferences between Categories D and E. Subject 23 has one usage of each category, so he changed his approach every time he started a problem.

SUBJECT	A	B	C	D	E	
21	0	0	5	0	0	
22	0	0	0	3	2	
23	1	1	1	1	1	
24	5	0	0	0	0	

Gross inspection suggests that Subject 21 is perfectly consistent, while Subject 23 is perfectly inconsistent. Most subjects will not fall at these extremes. For these intervening cases, it is convenient to adopt a "poker hand" type of consistency standard. Thus, four uses of one method and one use of some other method is more consistent than three uses of one method and two uses of another. Similarly, the three-one-one case shows more stability than two-two-one. Notice here that one category is just as good as another insofar as consistency is concerned--five consecutive uses of category A are entirely equivalent to five uses of category C; Subjects 21 and 24 are equally consistent. Our broad question of consistency can now be restated: Are the frequencies with which the men use the categories--the frequencies in this type of a matrix--significantly different from chance frequencies?

The theoretical probability distribution for such an experimental situation is difficult to specify because of the many interdependencies involved between successive trials by the same subjects. Accordingly, it was decided to manufacture, under the

experimental conditions, a sampling distribution which could be used for evaluating empirical frequencies.

The first step taken was to designate a consistency parameter. In the illustrative data matrix above, the rows are the basic elements in determining consistency. Though several different indices were considered, the row sum of squares was taken as the indicator of consistency. Obviously the sum of squares increases as the man's consistency increases: Subject 21 has a row sum of squares of 25, while subject 23 has only 5. Thus, the sum of squares index varies between 5 and 25: for the four-one case, it will be 17, for the three-two case it will be 13, and so on. For any given batch of data, the mean row sum of squares for the group was taken as the overall statistic.

The next step was to generate the sampling distribution by means of a large digital computer. This involved a sampling scheme wherein random numbers were accumulated according to the experimental conditions under which the technicians worked. In effect, an empty frequency matrix was set up in the computer's memory. Each "subject" had five "bins" to receive the category designations. A population of the category letters then was established, and a random selector was instructed to sample from this population and insert the letter designation in the subject's bin. To follow this through, consider Subject No. 1. At the appropriate instant, the random selector "finds" a category B in the population of letters and transmits that information to Subject 1's B storage. The next time around, another letter selection is made and sent down to Subject 1's storage. Of

course, it may or may not be another letter B. This process continues until Subject 1's storage bins accumulate a total of exactly five letters. When the bins for each subject have five entries in them, it is a simple matter for the computer to read out the frequencies, square and sum them, and average the sum of squares over all subjects. The final instruction to the computer is to print the mean row sum of squares, wipe out all the cells, and start another sampling cycle.

When a system like this is properly programmed, the accumulation, read out, and printing of the mean row sums of squares is very rapid. In two hours, for example, it is possible to generate 2,500 mean sums of squares when each mean is based on 70 cases. These 2,500 means, based on a total of 175,000 individual elements, comprise a quite adequate population distribution.

There are alternative ways that one can establish the pool of letter designations that the random selector operates upon. For example, we can set in equal number of A's, B's, C's, D's and E's, so that each letter has an equal chance of being chosen and sent to a memory element. This is the way to proceed when no advance information concerning category frequencies is at hand, or when the most general test is desired. Or we can put in A's, B's, C's and so on according to the actual popularity of the respective categories in the data being evaluated. Whichever course one adopts, the procedure and the type of end product remain essentially the same. It is clear, of course, that the particularized distribution—the one where probabilities of occurrence match those in the data—provides the preferred test.

Once the population distribution is available, its mean and variance can be determined, and the mean sum of squares from the empirical data compared with that of the population by conventional methods. The mean sum of squares population, for our data, is sufficiently normal for normal curve deviates to be employed in significance tests.

We can illustrate the application of the derived population distribution to some of our data. The average population sum of squares was 10.24 with a standard deviation of .48. The empirical average sum of squares from 70 technicians was 16.45, which was thirteen standard deviations or so above the population mean. This particular comparison was based on a reference distribution with the same proportion of cases in each category as the data sample. On these data, one can indeed assert that the men have stable preferences with respect to starting patterns.

In summary, an illustration is given that computer equipment can be programmed to generate reference distributions, provided that the restrictions on the data can be specified definitely enough. For some types of data, at least, the accumulation of values is very rapid, and may provide the only immediate way of making significance tests.

APPENDIX C

Supplementary Tables

Appendix Table A

Relationship Between Use of IAS Categories 1 Through 5 and Success on a Problem

Category	No. of Successful Performances	No. of Unsuccessful Performances	Total
1. Half-split	44	19	63
2. Middle-to-trouble	42	19	61
3. Loudspeaker-to-antenna	85	56	141
4. Antenna-to-loudspeaker	12	7	19
5. Unsystematic	28	38	66
Total	211	139	350

Chi square = 13.17, significant at .05 level for 4 df.

Note: Inter-category differences in proportion of success are mainly due to Category 5, and are no longer statistically significant when it is excluded. Comparing Unsystematic performances with performances in all other categories, the proportion of successes is statistically different. (Chi square = 10.84, significant at .001 level for 1 df.)

Appendix Table B

Analysis of Variance of Performance Lengths
Classified According to IAS Categories 1 through 5

Source	SS	df	MS	F
Between IAS Categories	3402.43	4	850.60	4.02**
Within IAS Categories	29366.37	139	211.26	
Total	32768.80	143		

**Significant at .01 level

Appendix Table C

Differences Between Performance Lengths Classified
According to IAS Categories 1 through 5
(t-ratios of differences between means)

	Category	N	1	2	3	4	5
1.	Half-split	63					
2.	Middle-to-trouble	61	1.17				
3.	Loudspeaker-to-antenna	141	2.08*	0.87			
4.	Antenna-to-loudspeaker	19	2.58**	1.77	1.36		
5.	Unsystematic	66	3.83**	2.60**	1.94	0.13	

* Significant at .05 level
**Significant at .01 level

Appendix Table D

Length of Initial Localizing Sequence

Data Group	Number of Actions	
	Mean	Standard Deviation
Job Sample Radio	11.72	9.83
Job Sample Radar	9.95	8.88
MASTS Radio	7.76	4.58
MASTS Radar	7.29	5.53
AUTOMASTS Radio	7.23	4.09
AUTOMASTS Radar	6.41	4.80

Appendix Table E

Relationship Between Amount of Prediagnostic Activity and Order in the Problem Series

Job Sample Data

Number of Prediagnostic Checks	Order of Administration 1	2	3	4	5	6	Total
More than Median Number	10	12	8	12	15	10	67
Fewer than Median Number	14	12	15	12	9	13	75
Total	24	24	23	24	24	23	142

Chi square = 4.09, not significant for 5 df.

MASTS Data

Number of Prediagnostic Checks	Order of Administration 1	2	3	4	5	6	Total
More than Median Number	15	10	16	14	15	11	81
Fewer than Median Number	15	18	17	16	18	17	101
Total	30	28	33	30	33	28	182

Chi square = 1.82, not significant for 5 df.

Note: Cell entries are number of performances. For each problem order, the performances are split at the integer nearest the grand median of the number of prediagnostic checks.

Appendix Table F

Breakdown of Replacements According to Component Types

	Radio		Radar	
	Number	Percentage of Total	Number	Percentage of Total
Tubes	268	30	343	25
Resistors	176	20	384	28
Capacitors	276	31	380	28
Connectors	172	19	257	19
Coils and Transformers	1	0	2	0
Total	893	100	1366	100

Appendix Table G

Breakdown of Radio Replacements According to Stage Location

Stage	Number of Replacements	Percentage of Total Replacements
RF Amplifier	33	4
RF Converter	119	14
IF Amplifier	249	28
Detector	369	41
Audio Voltage Amplifier	62	7
Phase Splitter	10	1
Audio Power Amplifier	39	4
Power Supply	12	1
Total	893	100

Appendix Table H

Breakdown of Radar Replacements According to Stage Location

Stage	Number of Replacements	Percentage of Total Replacements
Single Swing Blocking Oscillator	97	7
Diode Step Counter	339	25
Trigger Blocking Oscillator	214	16
Cathode Follower	64	4
One Shot Multivibrator	236	17
Saw Tooth Generator	216	16
Oscilloscope	148	11
Voltage Tripler	11	1
Bias Supply	2	0
Power Supply	39	3
Total	1366	100

Appendix Table I

Analysis of Variance Tables for Inter-Problem Action Rates of ETs in Homogeneous Format-Gear Groups

Data Group	Source	Sum of Squares	df	Variance Estimate
Job Sample Radio	Subjects	117.60	23	5.11
	Problems	39.97	5	7.99
	Error	93.09	115	.81
	Total	250.66	143	
Job Sample Radar	Subjects	55.07	23	2.39
	Problems	8.19	5	1.64
	Error	40.47	115	.35
	Total	103.73	143	
MASTS Radio	Subjects	59.81	23	2.60
	Problems	9.91	5	1.98
	Error	79.83	115	.69
	Total	149.56	143	
MASTS Radar	Subjects	77.97	23	3.39
	Problems	22.15	5	4.43
	Error	54.38	115	.47
	Total	154.50	143	
AUTOMASTS Radio	Subjects	144.46	44	3.28
	Problems	6.47	4	1.62
	Error	110.76	176	.63
	Total	261.69	224	
AUTOMASTS Radar	Subjects	151.78	44	3.45
	Problems	24.98	4	6.24
	Error	63.09	176	.36
	Total	239.85	224	

Formulas for consistency coefficients shown in Table 16 on page 142:

All possible pairs of problems: $$r_{11} = \frac{V_s - V_e}{V_s + (k-1)\,V_e}$$

All problems combined: $$r_{kk} = \frac{V_s - V_e}{V_s}$$

V_s = subject variance
V_e = error variance
k = number of problems